Hedging Futures with Spectral Risk Measures

This version: March 4, 2021

Abstract

We investigate different methods of hedging cryptocurrencies with Bitcoin futures. The introduction of derivatives on Bitcoin, in particular the launch of futures contracts on CME in December 2017, allows for hedging exposures on Bitcoin and cryptocurrencies in general. Because of volatility swings and jumps in Bitcoin prices, the traditional variance-based approach to obtain the hedge ratios is infeasible. The approach is therefore generalised to various risk measures, such as value-at-risk ,expected shortfall and spectral risk measures, and to different copulas for capturing the dependency between spot and future returns, such as the Gaussian, Student-t, NIG and Archimedean copulas. Various measures of hedge effectiveness in out-of-sample tests give insights in the practice of hedging Bitcoin and the CRIX index, a cryptocurrency index. This is joint work with Meng-Jou Lu (Asian Institute of Digital Finance Credit Research Initiative, National University of Singapore, Singapore) and Francis Liu (Berlin School of Economics and Law, Humboldt University Berlin).

JEL classification:
Keywords: Portfolio Selection, Spectral Risk Measurement, Coherent Risk

1 Introduction

Cryptocurrencies are a growing asset class. Bitcoin was the first cryptocurrency created in 2009 using a scheme proposed by Nakamoto (2019). Hedging is an important measure for investors to resist extreme risks and improve their profits. The hedge ratio for futures is the appropriate size futures contracts which should be held in short position in order to creak an opposite position. The determination of the fair number of futures is of course the difficulty in this hedging task. In this paper, we investigate different methods of hedging cryptocurrencies with Bitcoin futures. The approach is therefore generalised to various risk measures, such as the variance of the returns, value-at-risk (VaR), expected shortfall and spectral risk measures. The minimum variance hedge ratio is well known as a tool to obtain the optimal hedge ratio. However, it doesn't consider the investor's risk attitude. It is therefore important to describe the investor's behaviour when they choose to hedge the risk from spot market. In such an idealized stochastic framework downside risk, as determined by quantiles or VaR, the standard deviations (or variance) is all to know in order to hedge such positions. In realistic financial data scenarios though one cannot rely on only 2nd order moment calculations in order to minimize downside risk. The VaR as a sole risk measure has two disadvantages. First, it reflects only tail probability and not tail loss, and next it is not a coherent risk measure a very natural property that says that pooling will reduce risk.

This insight opens a new path of optimizing the hedge ratio by employing special risk measures, SRMs as financial risk measures. This paper expects that investors are more risk averse who might choose to accept with a low but guaranteed payment, rather than taking high risk of losing money but have high expected returns. In doing so we follow Barbi and Romagnoli (2014a) by applying exponential SRMs to determine the hedge ratio. By investigating the relationship of investors' utility function and the optimal hedge ratio, we demonstrate the relationship of the investors' risk aversion and SRMs, see Brandtner and Kürsten (2015).

SRM is a weighted average of the quantities of a loss distribution, the weights of which depend on the investor's risk-aversion. In other words, the risk estimation is directly related to the user's utility function. Popular examples are the exponential SRM and power SRM introduced by Dowd *et al.* (2008). Even though they reveal that SRM have some properties which cause problems when applying to practical risk management, they show that exponential utility function might be plausible in some circumstances under weak conditions Bühlmann, 1980. However, it still causes some problems to capture the behaviors of investors when the value of absolute risk aversion (ARA) parameter beyond a threshold Markowitz, 2014. If the relative risk aversion coefficients (RRA) are less than 1, Dowd *et al.* (2008) address that the weighting of lower risk-averse is higher than the higher risk-averse as the loss of portfolio increases. On the other hand, the power SRM proposed by Dowd *et al.* (2008) when the relative risk aversion coefficients (RRA) are larger than 1, has also proper features to give a higher weight as loss increase. Note that the selection of the utility function and the value of risk aversion parameter would be the matters of solving specific financial problem. By contrast, the estimation of the VaR and the ES are conditional on the confidence level which is not easy to determine. Since SRM is capable of reflecting the investor's attitude toward risk and has been applied to various fields of financial decision making, this paper apply to the determination of the optimal hedge ratio. It is important for the hedger who should choose a proper value for the hedge ratio in order to minimize the risk of portfolio.

A joint distribution of spot and futures has been specified in terms of a copula function to embody the tail behaviors of the spot and the futures Barbi and Romagnoli, 2014a. Copulae enable us to build the flexible multivariate distributions of dependence structure. This paper conducts four types of copulae (Gaussian, t, Frank, Clayton, and Normal Inverse Gaussian) to derive the copula representation of quantities to reach copula-based SRM of the hedged portfolio. It should be noted that the Clayton copula can be also used to construct the joint distribution with right tail dependence. Frank copula is symmetric and appropriate for data that exhibit weak or no tail dependence. Normal Inverse Gaussian (NIG) copula is a flexible system of joint distribution that includes fat-tailed and skewed distributions. However, there is still no evidence yet for selecting an exclusive copula in applications of hedging.

An optimal hedge ratio represents the investor's subjective marginal rate of substitution between risk and return. Cecchetti *et al.* (1988) found that the optimal hedge ratios increases when an investor with a greater risk aversion by maximizing the expected value of the logarithm of wealth. It is understandable if a investor's attitude is more risk-averse, they will increase the position of futures contracts to hedging the uncertain risk which they may take in the future. On the contrary, Brandtner and Kürsten (2015) address that the theoretical result predictions for the subset of exponential and power SRMs are not reasonable but may be counter-intuitive if the corresponding parameter of risk aversion is large enough. Different from Brandtner and Kürsten (2015), we consider the joint distribution of financial assets to choose the optimal hedge ratio by minimizing SRM. However, the empirical result shows the direction of optimal hedge ratio is increasing as the parameters represents the investors' attitude increases.

The remainder of the article is organized as follows. Section 2 methodology. Section 3 data, and Section 4 empirical result. Section 5 concludes.

2 Methodology

Following Barbi and Romagnoli (2014b), we consider the problem of optimal hedge ratios by extending the commonly known minimum variance hedge ratio to more general risk measures and dependence structures.

Hedge portfolio: $R_t^h = R_t^S - hR_t^F$, involving returns of spot and future contract and where h is the hedge ratio.

The optimal hedge ratio is

$$h^* = \underset{h}{\operatorname{argmin}}\, \rho_\phi(s, h), \tag{1}$$

for given confidence level $1 - s$ (if applicable, e.g. in the case of VaR, ES), where ρ_ϕ is a spectral risk measure with weighting function ϕ (see below). In other words, our task is to search for the optimal h which can minimize a particular risk measure. We call the risk measure being used in search process of h^* risk reduction objective. This naming is to differentiate the risk objective and risk outcome. One can see from result section that the h^* which minimize a particular risk measure in training does not necessarily minimize the risk measure in testing data. For example in table **??**, the best performing

risk reduction objective to reduce out-of-sample Value-at-Risk 99% is exponential risk measure $k = 10$.

The distribution function of R^h can be expressed in terms of the copula and the marginal distributions as Proposition 1 result shows (this is a corrected version of Corollary 2.1 of (Barbi and Romagnoli, 2014b)). For practical applications, it is numerically faster and more stable to use additional information about the specific copula and marginal distributions. We therefore derive semi-analytic formulas for a number of special cases, such as the Gaussian-, Student t-, normal inverse Gaussian (NIG) and Archimedean copulas in Section **??**.

Proposition 1 *Let R^S and R^F be two real-valued random variables on the same probability space $(\Omega, \mathcal{A}, p)$ with corresponding absolutely continuous copula $C_{R^S,R^F}(w, \lambda)$ and continuous marginals F_{R^S} and F_{R^F}. Then, the distribution of of R^h is given by*

$$F_{R^h}(x) = 1 - \int_0^1 D_1 C_{R^S,R^F}\left(u, F_{R^F}\left(\frac{F_{R^S}^{-1}(u) - x}{h}\right)\right) \mathrm{d}u. \tag{2}$$

Here $D_1 C(u, v) = \dfrac{\partial}{\partial u} C(u, v)$, which is easily shown to fulfil, see e.g. Equation (5.15) of (McNeil *et al.*, 2005):[1]

$$D_1 C_{X,Y}(F_X(x), F_Y(y)) = \mathbf{P}(Y \le y | X = x). \tag{3}$$

Proof. Using the identity (3) gives

$$\begin{aligned} F_{R^h}(x) &= \mathbf{P}(R^s - hR^F \le x) = \mathbb{E}\left[\mathbf{P}\left(R^F \ge \frac{R^S - x}{h} \Big| R^S\right)\right] \\ &= 1 - \mathbb{E}\left[\mathbf{P}\left(R^F \le \frac{R^S - x}{h} \Big| R^S\right)\right] = 1 - \int_0^1 D_1 C_{R^S,R^F}\left(u, F_{R^F}\left(\frac{F_{R^S}^{(-1)}(u) - x}{h}\right)\right) \mathrm{d}u. \end{aligned}$$

■

In addition to Barbi and Romagnoli (2014b) we propose an expression for the density of R^h

Proposition 2 *With the same setting of the above proposition, the density of r^h can be written as*

$$f_{R^h}(y) = \left|\frac{1}{h}\right| \int_0^1 c_{R^S,R^F}\left[u, F_{R^F}\left\{\frac{F_{R^S}^{-1}(u) - y}{h}\right\}\right] \cdot f_{R^F}\left\{\frac{F_{R^S}^{-1}(u) - y}{h}\right\} du \tag{4}$$

, or

$$f_{R^h}(y) = \int_0^1 c_{R^S,R^F}\left[u, F_{R^S}\left\{y + hF_{R^F}^{-1}(u)\right\}\right] \cdot f_{R^S}\left\{y + hF_{R^F}^{-1}(u)\right\} du. \tag{5}$$

[1] Let $F_X(x) = u$, $F_Y(y) = v$. Then, formally,

$$\begin{aligned} \frac{\partial}{\partial F_X(x)} C(F_X(x), F_Y(y)) &= \frac{\partial}{\partial F_X(x)} \mathbf{P}(U \le F_X(x), V \le F_Y(y)) = \mathbf{P}(U \in \mathrm{d}F_X(x), V \le F_Y(y)) \\ &= \mathbf{P}(V \le F_Y(y) | U = F_X(x)) \cdot \mathbf{P}(U \in \mathrm{d}F_X(x)) = \mathbf{P}(Y \le y | X = x) \cdot \mathbf{P}(U \in \mathrm{d}u) \\ &= \mathbf{P}(Y \le y | X = x). \end{aligned}$$

The two expressions are equivalent. One can use any of them to get the density of R^h. Notice that the density of R^h in the above proposition is readily accessible as long as one have the copula density and the marginal densities. The proof and a generic expression can be found in the appendix.

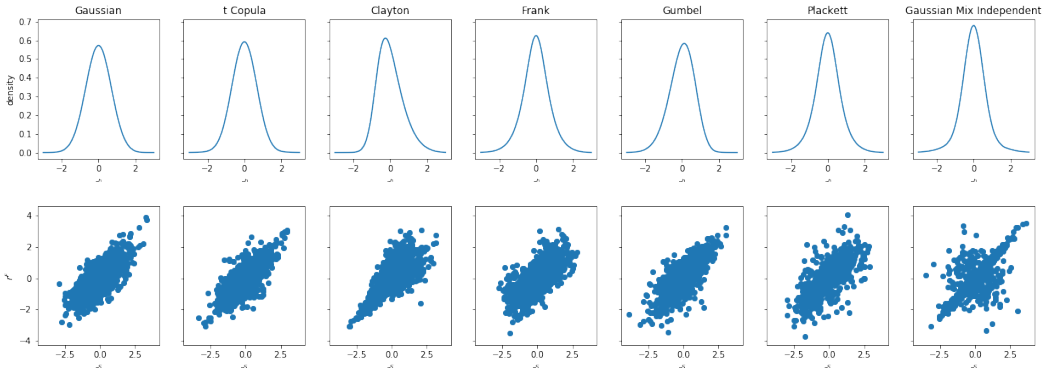

Figure 1: Upper Panel: Density of $Z = X - hY$ of different copulas with $X, Y \sim N(0, 1)$, 0.75 Spearman's rho between X and Y, and $h = 0.5$; Lower Panel: Scatter plot of samples from copulas. This illustration shows how dependence structure modelled by different copulas affects the density of the linear combination of margins. Notice that the Z modelled by the asymmetric copulas, namely the Clayton and Gumbel copulas, are skewed to right and left respectively.

2.1 Spectral Risk Measures

Spectral Risk Measures takes a form of

$$M_w(X) = -\int_0^1 w(p) q_s(X) ds \tag{6}$$

where $w(p)$ is a weighting function defined over the full range of cumulative probabilities $p \in [0, 1]$. M_w is a coherent measure if and only if w satisfies,

- Nonnegativity: $w(p) \geq 0$.
- Increasingness: $w'(p) \geq 0$.
- Normalisation: $\int_0^1 w(p)dp = 1$.

The first property requires that the weights are non-negative, and the second property is intended to reflect user risk aversion. The third one requires that the probability-weighted weights should sum to 1.

- Strict increasingness: $w'(p) > 0$.

Note that VaR and ES are included to spectral risk measure as special cases. The weighting function of VaR is a Dirac delta function which gives the outcome an infinite weight and the others a zero weight. On the other hand, the ES gives all tail quantiles the same weight. Both of them are not a suitable weight function for capturing investor's risk attitudes.

By setting a 'well-behaved' risk-aversion function which indicates the weights will rise more rapidly when the degree of risk aversion is higher, we investigate the behaviors of the users in terms of different weight function when they determine the hedge ratios.

We also consider wildly used risk measure is Value at Risk, VaR, a quantile of the portfolio loss distribution (Jürgen *et al.*, 2011)

$$q_\alpha(X) = F_X^{-1}(\alpha), \quad \alpha \in (0,1). \tag{7}$$

For any random variable X, and its cumulative distribution function F_X is well defined. Due to the inconsistency of coherent risk, the use of expected shortfall has been discussed intensively in finance and risk management (Jürgen *et al.*, 2011). Expected shortfall (ES) measures are expressed as

$$\mathrm{ES}_\alpha(X) = \frac{1}{1-\alpha}\int_\alpha^1 q_s(X)ds \tag{8}$$

2.2 Two Risk Spectra

Recognising the importance of the weighting function, we investigate different utility functions, $U(x)$ defined over outcomes x. Consider the exponential utility and power utility, where the investor's coefficient of absolute risk aversion is $k(x) = -\frac{U''(x)}{U'(x)}$ and his relative risk-aversion is $\gamma(x) = -\frac{xU''(x)}{U'(x)}$. This allows us to transfer the utility function to a weighting function as in Dowd *et al.* (2008).

2.2.1 Exponential Spectral Risk Measure

The exponential SRM is specified by only one risk parameter k. To obtain the risk spectrum, we set $w(p) = \lambda e^{-k(1-p)}$ and $\lambda = \frac{k}{1-e^{-k}}$. Then, the risk spectrum and its antiderivative are:

$$w(p) = \frac{ke^{-k(1-p)}}{1-e^{-k}}, \quad \text{and} \quad W(p) = -\frac{1-e^{-k(1-p)}}{1-e^{-k}} \tag{9}$$

where $k \in (0,\infty), \quad p \in [0,1]$. This function depends on only one k. Figure 2 shows the exponential risk spectrum and its antiderivative for $k = 1$, and 2. By substituting into (6), the exponential SRM is,

$$M_w(X) = \int_0^1 \frac{ke^{-k(1-p)}}{1-e^{-k}} F^{-1}(p)dp \tag{10}$$

2.2.2 Power Spectral Risk Measure

The power weighting function only has one parameter, γ, which leads to $w(p) = \frac{\lambda(1-p)^{\gamma-1}}{1-\gamma}$ as $0 < \gamma < 1$. Then, by setting $\lambda = \gamma(1-\gamma)$, the risk spectrum and its antiderivative are,

$$w(p) = \gamma(1-p)^{\gamma-1}, \quad \text{and} \quad W(p) = -(1-p)^\gamma \tag{11}$$

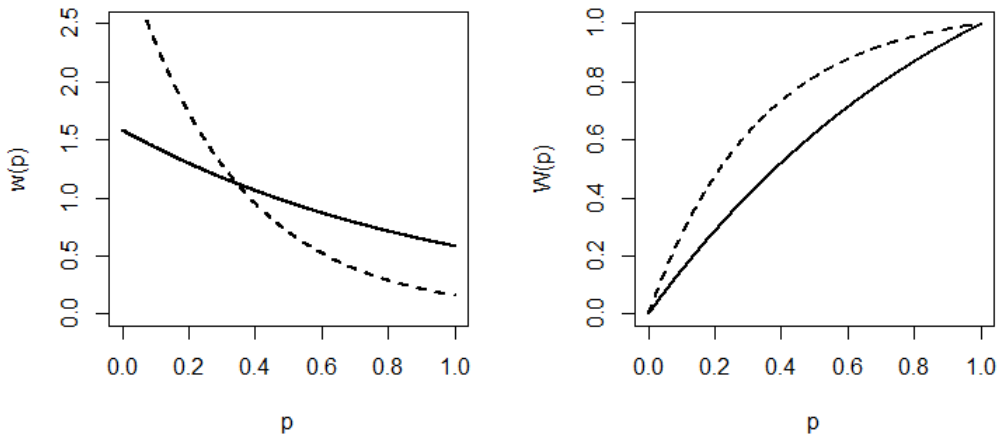

Figure 2: Exponential SRMs for $k = 1$ (dashed) and $k = 2$ (solid).

Plugging the weighting function to (6), the power SRM is obtained,

$$M_w(X) = \int_0^1 \gamma(1-p)^{\gamma-1} F^{-1}(p)dp \tag{12}$$

For the case of $\gamma > 1$, we have $w(p) = \frac{\lambda p^{\gamma-1}}{1-\gamma}$ with $\lambda = \gamma(1-\gamma)$. The risk spectrum is written as,

$$w(p) = \gamma p^{\gamma-1} \tag{13}$$

The Power SRM then becomes,

$$M_w(X) = \int_0^1 \gamma p^{\gamma-1} F^{-1}(p)dp \tag{14}$$

2.3 Copulas

We test different copulas' ability to model crypto-currency data, they include Gaussian-, t-, Frank-, Gumbel-, Clayton-, Plackett-, mixture, and factor copula. We will go through the copulas in this section.

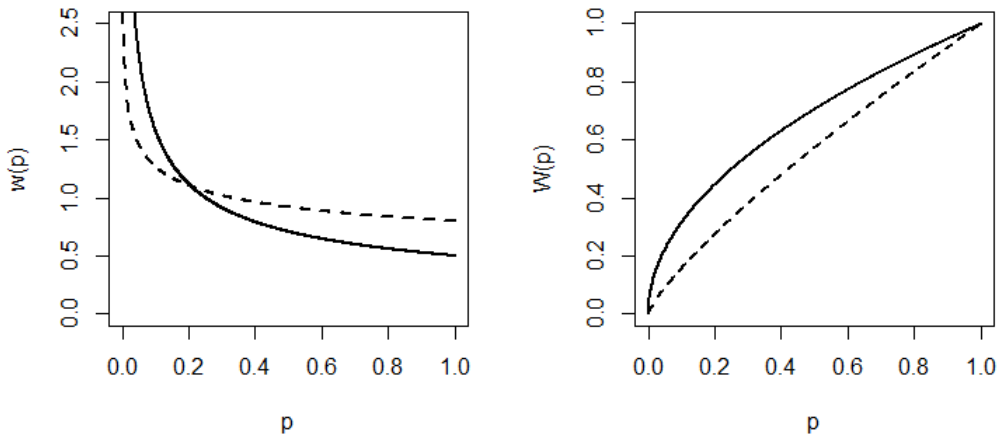

Figure 3: Power SRMs for $\gamma = 0.5$ (solid) and $\gamma = 0.8$ (dashed).

2.3.1 Ellpitical Copulas

Elliptical copulas are copulas of elliptical distributions. Gaussian copula is the copula associated with multivariate normal distribution. The Gaussian copula has a form

$$\boldsymbol{C}(u, v) = \Phi_{2,\rho}\{\Phi^{-1}(u), \Phi^{-1}(v)\} \tag{15}$$

$$= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right) dsdt \tag{16}$$

where $\Phi_{2,\rho}$ is the cdf of bivariate Normal distribution with zero mean, unit variance, and correlation ρ, and Φ^{-1} is quantile function univariate standard normal distribution. The Gaussian copula density is

$$\boldsymbol{c}_\rho(u, v) = \frac{\varphi_{2,\rho}\{\Phi^{-1}(u), \Phi^{-1}(v)\}}{\varphi\{\Phi^{-1}(u)\} \cdot \varphi\{\Phi^{-1}(v)\}} \tag{17}$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{u^2 - 2\rho uv + v^2}{2(1-\rho^2)}\right), \tag{18}$$

where $\phi_{2,\rho}(\cdot)$ is the density of bivariate Normal distribution with zero mean, unit variance, and correlation ρ, and, $\phi(\cdot)$ the density of standard normal distribution.

The Kendall's τ_K and Spearman's ρ_S of a bivariate Gaussian Copula are

$$\tau_K(\rho) = \frac{2}{\pi} \arcsin \rho \tag{19}$$

$$\rho_S(\rho) = \frac{6}{\pi} \arcsin \frac{\rho}{2} \tag{20}$$

t-copula is associated with multivariate t distribution. The t Copula takes a form

$$\boldsymbol{C}(u,v) = \boldsymbol{T}_{2,\rho,\nu}\{T_\nu^{-1}(u), T_\nu^{-1}(v)\} \tag{21}$$

$$= \int_{-\infty}^{T_\nu^{-1}(u)} \int_{-\infty}^{T_\nu^{-1}(v)} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\pi\nu\sqrt{1-\rho^2}} \tag{22}$$

$$\left(1 + \frac{s^2 - 2st\rho + t^2}{\nu}\right)^{-\frac{\nu+2}{2}} dsdt, \tag{23}$$

where $\boldsymbol{T}_{2,\rho,\nu}(\cdot,\cdot)$ denotes the cdf of bivariate t distribution with scale parameter ρ and degree of free ν, $T_\nu^{-1}(\cdot)$ is the quantile function of a standard t distribution with degree of freedom ρ.

The copula density is

$$\boldsymbol{c}(u,v) = \frac{\boldsymbol{t}_{2,\rho,\nu}\{T_\nu^{-1}(u), T_\nu^{-1}(v)\}}{t_\nu\{T_\nu^{-1}(u)\} \cdot t_\nu\{T_\nu^{-1}(v)\}}, \tag{24}$$

where $\boldsymbol{t}_{2,\rho,\nu}$ is the density of bivariate t distribution, and t_ν the density of standard t distribution.

Like all the other elliptical copula, t copula's Kendall's τ is same to that of Gaussian copula (Demarta and reference therein).

2.3.2 Archimedean Copula

Archimedean copula forms a large class of copulas with many convenient features.

In general, Archimedean copula takes a form

$$\boldsymbol{C}(u,v) = \psi^{-1}\{\psi(u), \psi(v)\}, \tag{25}$$

where $\psi : [0,1] \to [0,\infty)$ is a continuous, strictly decreasing and convex function such that $\psi(1) = 0$ for any permissible dependence parameter θ. ψ is also called generator. ψ^{-1} is the inverse the generator.

Frank copula (B3 in Joe) is a radial symmetric copula and does not have any tail dependence. It takes the form

$$\boldsymbol{C}_\theta(u,v) = \frac{1}{\log(\theta)} \log\left\{1 + \frac{(\theta^u - 1)(\theta^v - 1)}{\theta - 1}\right\} \tag{26}$$

where $\theta \in [0,\infty]$ is the dependency parameter. $\boldsymbol{C}_1 = \boldsymbol{M}$, $\boldsymbol{C}_1 = \boldsymbol{\Pi}$, and $\boldsymbol{C}_\infty = \boldsymbol{W}$.

The Copula density is

$$\boldsymbol{c}_\theta(u,v) = \frac{(\theta-1)\theta^{u+v}\log(\theta)}{\theta^{u+v} - \theta^u - \theta^v + \theta} \tag{27}$$

Frank copula has Kendall's τ and Spearman's ρ as follow:

$$\tau_K(\theta) = 1 - 4\frac{D_1\{-\log(\theta)\}}{\log(\theta)}, \tag{28}$$

and

$$\rho_S(\theta) = 1 - 12\frac{D_2\{-\log(\theta)\} - D_1\{\log(\theta)\}}{\log(\theta)}, \tag{29}$$

where D_1 and D_2 are the Debye function of order 1 and 2. Debye function is $D_n = \frac{n}{x^n}\int_0^x \frac{t^n}{e^t-1}dt$.

Gumbel copula (B6 in Joe) has upper tail dependence with the dependence parameter $\lambda^U = 2-2^{\frac{1}{\theta}}$ and displays no lower tail dependence.

$$\boldsymbol{C}_\theta(u,v) = \exp -\{(-\log(u))^\theta + (-\log(v))^\theta\}^{\frac{1}{\theta}}, \tag{30}$$

where $\theta \in [1,\infty)$ is the dependence parameter.While Gumbel copula cannot model perfect counter dependence (ref), $\boldsymbol{C}_1 = \boldsymbol{\Pi}$ models the independence, and $\lim_{\theta\to\infty} \boldsymbol{C}_\theta = \boldsymbol{W}$ models the perfect dependence.

$$\tau_K(\theta) = \frac{\theta-1}{\theta} \tag{31}$$

Clayton copula, opposite to Gumbel copula, generates lower tail dependence in a form $\lambda^L = 2^{-\frac{1}{\theta}}$, but generates no upper tail dependence.

Clayton copula takes a form

$$\boldsymbol{C}_\theta(u,v) = \left[\max\{u^{-\theta}+v^{-\theta}-1,0\}\right]^{-\frac{1}{\theta}}, \tag{32}$$

where $\theta \in (-\infty,\infty)$ is the dependency parameter. $\lim_{\theta\to-\infty} \boldsymbol{C}_\theta = \boldsymbol{M}$, $\boldsymbol{C}_0 = \boldsymbol{\Pi}$, and $\lim_{\theta\to\infty} \boldsymbol{C}_\theta = \boldsymbol{W}$.

Its Kendall's τ is

$$\tau_K(\theta) = \frac{\theta}{\theta+2}. \tag{33}$$

Plackett copula has an expression

$$\boldsymbol{C}_\theta(u,v) = \frac{1+(\theta-1)(u+v)}{2(\theta-1)} - \frac{\sqrt{\{1+(\theta-1)(u+v)\}^2 - 4uv\theta(\theta-1)}}{2(\theta-1)} \tag{34}$$

$$\rho_S(\theta) = \frac{\theta+1}{\theta-1} - \frac{2\theta\log\theta}{(\theta-1)^2} \tag{35}$$

We include Placket copula in our analysis as it possesses a special property, the cross-product ratio is equal to the dependence parameter

$$\frac{\mathbf{P}(U\le u, V\le v)\cdot\mathbf{P}(U>u,V>v)}{\mathbf{P}(U\le u,V>v)\cdot\mathbf{P}(U>u,V\le v)} \tag{36}$$

$$= \frac{\boldsymbol{C}_\theta(u,v)\{1-u-v+\boldsymbol{C}_\theta(u,v)\}}{\{u-\boldsymbol{C}_\theta(u,v)\}\{v-\boldsymbol{C}_\theta(u,v)} \tag{37}$$

$$= \theta. \tag{38}$$

That is, the dependence parameter is equal to the ratio between number of concordence pairs and number of discordence pairs of a bivariate random variable.

2.3.3 Mixture Copula

Mixture copula is a linear combination of copulas. It allows us to model the dependence structure in a more flexible manner.

For a 2-dimensional random variable $\boldsymbol{X} = (X_1, X_2)^\intercal$, its distribution can be written as linear combination K copulas

$$\mathbf{P}(X_1 \leq x_1, X_2 \leq x_2) = \sum_{k=1}^{K} p^k \cdot \boldsymbol{C}^{(k)}\{F_{X_1}^{(k)}(x_1; \boldsymbol{\gamma}_1^{(k)}), F_{X_2}^{(k)}(x_2; \boldsymbol{\gamma}_2^{(k)}); \boldsymbol{\theta^{(k)}}\} \tag{39}$$

where $p^{(k)} \in [0, 1]$ is the weight of each component, $\boldsymbol{\gamma}^{(k)}$ is the parameter of the marginal distribution in the k^{th} component, and $\boldsymbol{\theta^{(k)}}$ is the dependence parameter of the k^{th} component. We also restrict the weight so that $\sum_{k=1}^{K} p^{(k)} = 1$. Analysis of mixture copula with higher dimension can be found in Vrac et. al. (2011).

We deploy a simplified version of the above representation by assuming the maringals of $\boldsymbol{X}$ are not mixture. By Sklar's theorem we write

$$\boldsymbol{C}(u, v) = \sum_{k=1}^{K} p^{(k)} \cdot \boldsymbol{C}^{(k)}\{F_{X_1}^{-1}(u), F_{X_2}^{-1}(v); \boldsymbol{\theta^{(k)}}\}. \tag{40}$$

The copula density is again a linear combination of copula density

$$\boldsymbol{c}(u, v) = \sum_{k=1}^{K} p^{(k)} \cdot \boldsymbol{c}^{(k)}\{F_{X_1}^{-1}(u), F_{X_2}^{-1}(v); \boldsymbol{\theta^{(k)}}\}. \tag{41}$$

While Kendall's τ of mixture copula is not known in close form, the Spearman's ρ is

Proposition 3 *Let $\rho_S^{(k)}$ be the Spearman's ρ of the k^{th} component and $\sum_{k=1}^{K} p^{(k)} = 1$ holds, the Spearman's ρ of a mixture copula is*

$$\rho_S = \sum_{k=1}^{K} p^{(k)} \cdot \rho_S^{(k)} \tag{42}$$

Proof. Spearman's ρ is defined as (Nelsen)

$$\rho_S = 12 \int_{\mathbb{I}^2} \boldsymbol{C}(s, t) ds dt - 3. \tag{43}$$

Rewrite the mixture copula into sumation of components

$$\rho_S = 12 \int_{\mathbb{I}^2} \sum_{k=1}^{K} p^{(k)} \cdot \boldsymbol{C}^{(k)}(s, t) ds dt - 3. \tag{44}$$

■

Example 4 *Frechet class can be seen as an example of mixture copula. It is a convex combinations of $\boldsymbol{W}$, $\boldsymbol{\Pi}$, and $\boldsymbol{M}$ (Nelsen)*

$$\boldsymbol{C}_{\alpha,\beta}(u, v) = \alpha \boldsymbol{M}(u, v) + (1 - \alpha - \beta)\boldsymbol{\Pi}(u, v) + \beta \boldsymbol{W}(u, v), \tag{45}$$

where α and β are the dependence parameters, with $\alpha, \beta \geq 0$ and $\alpha + \beta \leq 1$. Its Kendall's τ and Spearman's ρ are

$$\tau_K(\alpha, \beta) = \frac{(\alpha - \beta)(\alpha + \beta + 2)}{3} \tag{46}$$

, and

$$\rho_S(\alpha, \beta) = \alpha - \beta \tag{47}$$

We use a mixture of Gaussian and independent copula in our analysis. We write the copula

$$\boldsymbol{C}(u, v) = p \cdot \boldsymbol{C}^{\text{Gaussian}}(u, v) + (1 - p)(uv). \tag{48}$$

The corresponding copula density is

$$\boldsymbol{c}(u, v) = p \cdot \boldsymbol{c}^{\text{Gaussian}}(u, v) + (1 - p). \tag{49}$$

This mixture allows us to model how much "random noise" appear in the dependency structure. In this hedging exercise, the structure of the "random noise" is not of our concern nor we can hedge the noise by a two-asset portfolio. However, the proportion of the "random noise" does affect the distribution of R^h (see figure), so as the optimal hedging ratio h^* (see figure). One can consider the mixture copula as a handful tool for stress testing. Similar to this Gaussian mix Independent copula, t copula is also a two parameter copula allow us to model the noise, but its interpretation of parameters is not as intuitive as that of a mixture. The mixing variable p is the proportion of a manageable (hedgable?) Gaussian copula, while the remaining proportion $1 - p$ cannot be managed.

3 Estimation

3.1 Simulated Method of Moments

This method is suggested by Oh and Patton (2013). In this setting, rank correlation e.g. Spearman's ρ or Kendall's τ, and quantile dependence measures at different levels λ_q are calibrated against their empirical counterparts.

Spearman's rho, Kendall's tau, and quantile dependence of a pair (X, Y) with copula C are defined as

$$\rho_S = 12 \int \int_{I^2} C_{\boldsymbol{\theta}}(u, v) \, \mathrm{d}u \, \mathrm{d}v - 3 \tag{50}$$

$$\tau_K = 4\mathbb{E}[C_{\boldsymbol{\theta}}\{F_X(x), F_Y(y)\}] - 1, \tag{51}$$

$$\lambda_q = \begin{cases} \mathbf{P}(F_X(X) \leq q | F_Y(Y) \leq q) = \dfrac{C_{\boldsymbol{\theta}}(q, q)}{q}, & \text{if } q \in (0, 0.5], \\ \mathbf{P}(F_X(X) > q | F_Y(Y) > q) = \dfrac{1 - 2q + C_{\boldsymbol{\theta}}(q, q)}{1 - q}, & \text{if } q \in (0.5, 1). \end{cases} \tag{52}$$

The empirical counterparts are

$$\hat{\rho}_S = \frac{12}{n}\sum_{k=1}^{n}\hat{F}_X(x_k)\hat{F}_Y(y_k) - 3,$$

$$\hat{\tau}_K = \frac{4}{n}\sum_{k=1}^{n}\hat{C}\{\hat{F}_X(x_i), \hat{F}_X(y_i)\} - 1,$$

$$\hat{\lambda}_q = \begin{cases} \frac{1}{n}\sum_{k=1}^{n}\frac{\mathbf{1}_{\{\hat{F}_X(x_k)\leq q, \hat{F}_Y(y_k)\leq q\}}}{q}, & \text{if } q \in (0, 0.5], \\ \frac{1}{n}\sum_{k=1}^{n}\frac{\mathbf{1}_{\{\hat{F}_X(x_k)> q, \hat{F}_Y(y_k)> q\}}}{1-q}, & \text{if } q \in (0.5, 1). \end{cases},$$

where $\hat{F}(x) := \frac{1}{n}\sum_{k=1}^{n} 1_{\{x_i \leq x\}}$ and $\hat{C}(u, v) := \frac{1}{n}\sum_{k=1}^{n} 1_{\{u_i \leq u, v_i \leq v\}}$.

We denote $\tilde{\boldsymbol{m}}(\boldsymbol{\theta})$ be a m-dimensional vector of dependence measures according the the dependence parameters $\boldsymbol{\theta}$,and $\hat{\boldsymbol{m}}$ be the corresponding empirical counterpart. The difference between dependence measures and their counterpart is denoted by

$$\boldsymbol{g}(\boldsymbol{\theta}) = \hat{\boldsymbol{m}} - \tilde{\boldsymbol{m}}(\boldsymbol{\theta}).$$

The SMM estimator is

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmin}}\, \boldsymbol{g}(\boldsymbol{\theta})^{\intercal}\hat{\boldsymbol{W}}\boldsymbol{g}(\boldsymbol{\theta}),$$

where $\hat{W}$ is some positive definite weigh matrix.

In this work, we use $\tilde{\boldsymbol{m}}(\boldsymbol{\theta}) = (\rho_S, \lambda_{0.05}, \lambda_{0.1}, \lambda_{0.9}, \lambda_{0.95})^{\intercal}$ for calibration of Bitcoin price and CME Bitcoin future.

3.2 Maximum Likelihood Estimation

By Sklar's theorem, the joint density of a d-dimensional random variable $\boldsymbol{X}$ with sample size n can be written as

$$\boldsymbol{f}_{\boldsymbol{X}}(x_1, ..., x_d) = \boldsymbol{c}\{F_{X_1}(x_1), ..., F_{X_d}(x_d)\}\prod_{j=1}^{d} f_{X_i}(x_i). \tag{53}$$

We follow the treatment of MLE documented in section 10.1 of Joe (1997), namely the inference functions for margins or IFM method. The log-likelihood $\sum_{i=1}^{n} f_{\boldsymbol{X}}(X_{i,1}, ..., X_{i,d})$ can be decomposed into dependence part and marginal part,

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{n}\boldsymbol{c}\{F_{X_1}(x_{i,1}; \boldsymbol{\delta}_1), ..., F_{X_d}(x_{i,d}; \boldsymbol{\delta}_d); \boldsymbol{\gamma}\} + \sum_{i=1}^{n}\sum_{j=1}^{d} f_{X_j}(x_{i,j}; \boldsymbol{\delta}_j) \tag{54}$$

$$= L_C(\boldsymbol{\delta}_1, ..., \boldsymbol{\delta}_d, \boldsymbol{\gamma}) + \sum_{j=1}^{d} L_j(\boldsymbol{\delta}_j) \tag{55}$$

where $\boldsymbol{\delta}_j$ is the parameter of the j-th margin, $\boldsymbol{\gamma}$ is the parameter of the parametric copula, and $\boldsymbol{\theta} = (\boldsymbol{\delta}_1, ..., \boldsymbol{\delta}_d, \boldsymbol{\gamma})$.

Instead of searching the $\boldsymbol{\theta}$ is a high dimensional space, Joe (1997) suggests to search for $\hat{\boldsymbol{\delta}}_1, ..., \hat{\boldsymbol{\delta}}_d$ that maximize $L_1(\boldsymbol{\delta}_1), ..., L_d(\boldsymbol{\delta}_d)$, then search for $\hat{\boldsymbol{\gamma}}$ that maximize $L_C(\hat{\boldsymbol{\delta}}_1, ..., \hat{\boldsymbol{\delta}}_d, \boldsymbol{\gamma})$.

That is, under regularity conditions, $(\hat{\boldsymbol{\delta}}_1, ..., \hat{\boldsymbol{\delta}}_d, \hat{\boldsymbol{\gamma}})$ is the solution of

$$\left(\frac{\partial L_1}{\partial \boldsymbol{\delta}_1}, ..., \frac{\partial L_d}{\partial \boldsymbol{\delta}_d}, \frac{\partial L_C}{\partial \boldsymbol{\gamma}}\right) = \mathbf{0}. \tag{56}$$

However, the IFM requires making assumption to the distribution of of the margins. Genest *et al.* (1995) suggests to replace the estimation of marginals parameters estimation by non-parameteric estimation. Given non-parametric estimator $\hat{F}_i$ of the margins F_i, the estimator of the dependence parameters $\boldsymbol{\gamma}$ is

$$\hat{\boldsymbol{\gamma}} = \underset{\boldsymbol{\gamma}}{\operatorname{argmax}} \sum_{i=1}^{n} \boldsymbol{c}\{\hat{F}_{X_1}(x_{i,1}), ..., \hat{F}_{X_d}(x_{i,d}); \boldsymbol{\gamma}\}. \tag{57}$$

3.3 Comparison

Both the simulated method of moments and the maximum likelihood estimation are unbiased and proven to give good fits. The problem remain is which procedure is more suitable for hedging. Sample and fitted quantile dependence for Bitcoin and CME future.

The MM estimation perform just as we decided: match the upper and lower quantile dependence.

4 Data and margins

TODO:

- Data description
- KDE and bandwidth selection
- plots
- Data preprocessing, the 300-30 train-test split, moving window estimation etc

5 Results

We illustrate the results in three directions, hedging effectiveness, ability of hedging extreme negative events in R^S, and the stability of h^*.

- Hedging Effectiveness
  - Kick out Frank for its ineffectiveness; Alternative to a one-parameter symmetric Archimedean copula is Plackett;
  - Differences among combinations of copula and risk reduction objective are small;
  - None of the combination can escape from the structural break point (dependence of training is stronger then that of testing). (The bump in 25-26th Sept 2019)
  - The best performing RRO of a particular risk measure in out-of-sample R^h is not necessarily same, e.g. VaR 95% as RRO (with Gumbel copula) can generate the lowest out-of-sample ES 99%.

- Ability to hedge extreme events in R^S
  - The extreme events in R^S are well managed by the hedge. The magnitudes of loss in R^h is much smaller than that of R^h. (Visually seen from the time series of R^h)
  - None of the combination can escape from the structural break point (dependence of training is stronger then that of testing)
- Stability of h^*
  - Gumbel gives high h all the time; the extreme events are "hedged" ex-ante.
  - ES 99% and VaR 99% as risk reduction objective are too sensitive to extremes in training data; Large changes in h are suggested in response to extremes training data, while the testing data are less extreme;
  - ERM can be seen as a smoothed risk measure focusing in the lower tail of R^h; Less sensitive to rare events; Suggested.
-

5.1 Hedging Effectiveness

The hedging effectiveness(HE) is defined as

$$1 - \frac{\rho_\phi(R^h)}{\rho_\phi(R^S)}. \tag{58}$$

The hedging effectiveness is the reduction of portfolio risk. This way of evaluating of hedging performance is proposed by Ederington (1979) in the context of, at that time, hedging the newly introduced organized futures market. He evaluates the extent of variance reduction by introducing another asset. We measure the hedging effectiveness also in other risk measure mentioned in section 2.1, for example

$$1 - \frac{\mathrm{ES}_\alpha(R^h)}{\mathrm{ES}_\alpha(R^S)}. \tag{59}$$

The box-plots in figure 8 show the out-of-sample hedging effectiveness of different copulas under various risk reduction objectives across testing datasets. Observe that in most of the copulas perform well in most of the time. The average HE of copulas and risk reduction objectives is higher than 60% except for Frank-copula. However, the HEs vary a lot in different testing data. In some instances, the HE can be as low as 10%. This reflects the highly violate nature of cryptocurrencies: the optimal hedge ratio in the training data deviates from that of testing data. There is a large literature about structural break points and time changing dependence, to name a few Hafner and Manner (2012), Patton (2006), Creal *et al.* (2008), Engle (2002), and Giacomini *et al.* (2009). Manner and Reznikova (2012) gives a great survey about this issue. The discussion is out of the scope of this study.

Frank-copula, in general, is not a good choice to model financial data. Figure 4

Aside from the Frank-copula, the HEs of various combination of copula and risk reduction objective are very similar. This is an expected result as the portfolio consists only two assets. In addition to hedging effectiveness, we observe the out-of-sample returns of the hedged portfolio. Figure 6 tabulates the time series of out-of-sample returns of hedged portfolio under various copulas and risk reduction objectives.

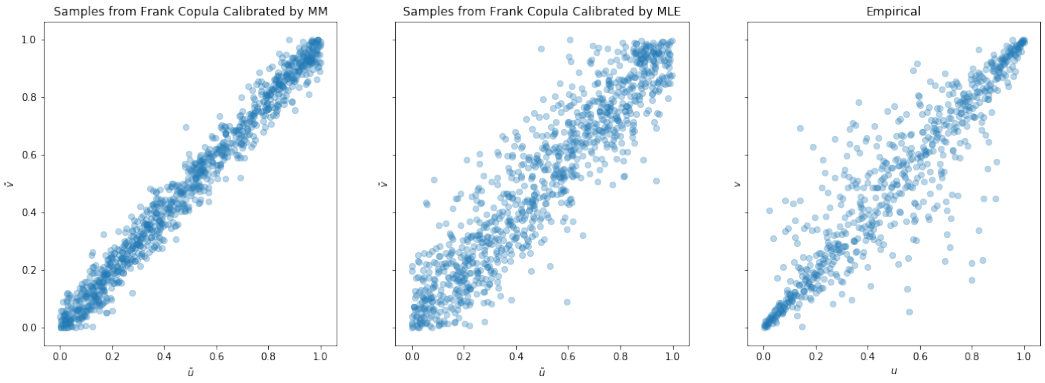

Figure 4: Comparison of Frank Copula Samples and Pseudo Observations of Bitcoin and CME Future Returns.

One can see all the combinations of copula and risk reduction objective generate a large fluctuation of returns in 25/09/2019 and 26/09/2019. This large fluctuation is due to dependence break.

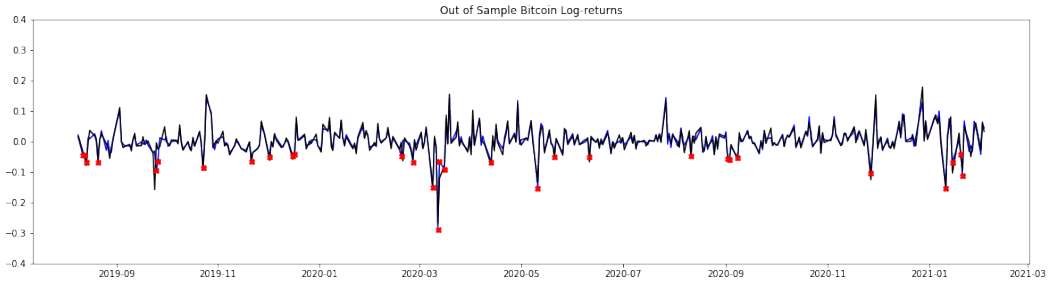

Figure 5: Out of Sample Log-return of Bitcoin.

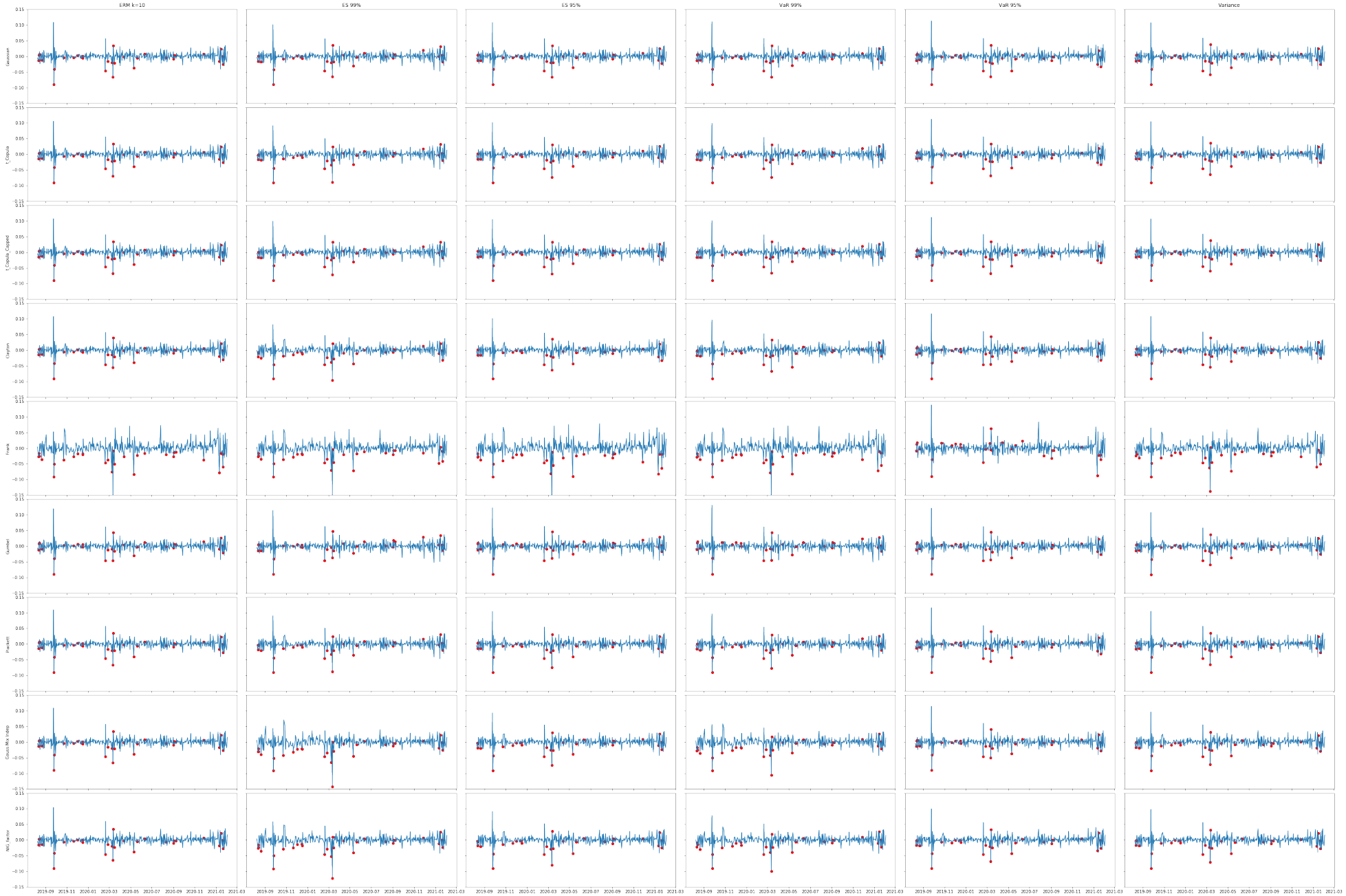

Figure 6: Out-of-Sample Returns of Hedged Portfolio of Copulas and Risk Reduction Objectives.

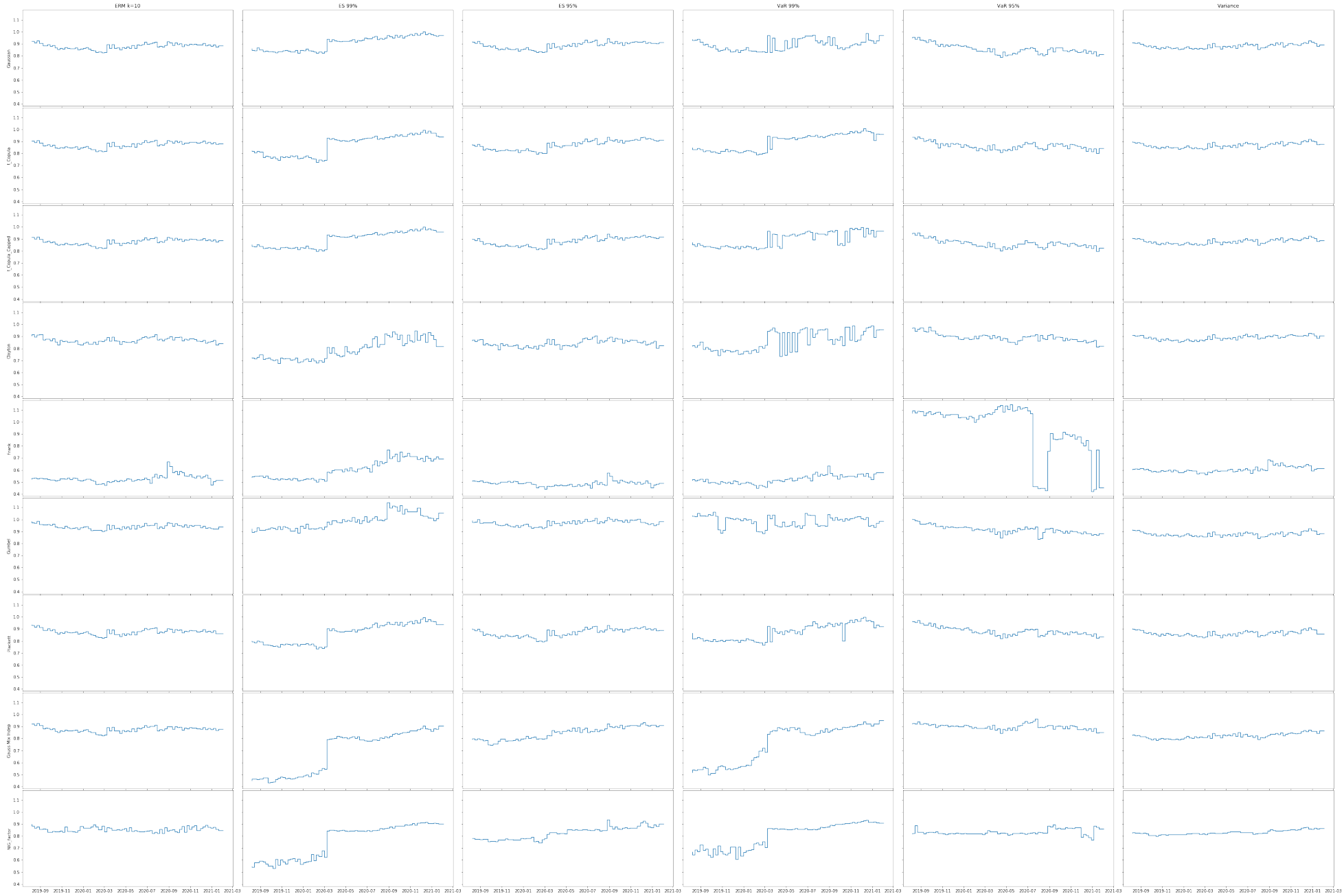

Figure 7: Optimal Hedge Ratio Obtained from Combinations of Copula and Risk Reduction Objective.

Figure 5 shows the time series of out-of-sample R^h using Gumbel copula with the objective of reducing variance. The red dots are the 30 most extreme negative returns in Bitcoin. In the figure, we can see the downside risk of Bitcoin is well managed by the hedging procedure with Gumbel copula. Most of the extreme losses of Bitcoin are greatly reduced by introducing the CME future in the hedged portfolio. Two exceptions are found in 25/09/2019 and 26/09/2019, where the CME future failed to follow the large drop in Bitcoin. (TODO: drop reason) One of the possible reason is that traders was performing rollover activities on 25-26/09/2019, which 27/09/2019 is the expiry day of the September future. Another reason for Gumbel fail of capturing the loss is dependence break. The Kendall's tau in the training data is 0.2 higher than that of the testing data. Other copulas suffer from the break as well.

5.2 Stability of h^*

We measure the stability of h^* by sum of absolute change

$$\sum_{t=1}^{T} |h_t - h_{t-1}|. \tag{60}$$

Adjustment of portfolio weights induces price slippage (ref) and transaction cost. From figure **??** we know the NIG factor copula with variance as risk reduction objective generates the smallest sum of absolute change in OHR.

	ERM k=10	ES 99%	ES 95%	VaR 99%	VaR 95%	Variance
Gaussian	0.019985	0.020802	0.020061	0.020230	0.019983	0.019757
t_Copula	0.020097	0.021698	0.020381	0.020966	0.020071	0.019890
t_Copula_Capped	0.020048	0.021018	0.020202	0.020554	0.020059	0.019792
Clayton	0.019519	0.021341	0.019789	0.021045	0.019389	0.019675
Frank	0.029234	0.026240	0.030770	0.029157	0.023085	0.025928
Gumbel	0.020014	0.021411	0.020511	0.021643	0.019557	0.019757
Plackett	0.020010	0.021531	0.020363	0.020870	0.019755	0.019909
Gauss Mix Indep	0.019949	0.025390	0.020454	0.023283	0.019667	0.020006
NIG_factor	0.019720	0.023425	0.020706	0.022039	0.019950	0.019999

Table 1: Exponential Risk Measure $k = 10$

	ERM k=10	ES 99%	ES 95%	VaR 99%	VaR 95%	Variance
Gaussian	0.061084	0.062405	0.061201	0.062148	0.061712	0.059310
t_Copula	0.062148	0.068702	0.063339	0.063964	0.062067	0.060735
t_Copula_Capped	0.061623	0.064114	0.062198	0.062466	0.062072	0.059676
Clayton	0.058495	0.069910	0.060812	0.064595	0.055962	0.058318
Frank	0.104185	0.096795	0.108713	0.105070	0.068457	0.091321
Gumbel	0.056513	0.059574	0.056035	0.058162	0.055492	0.059525
Plackett	0.061167	0.068027	0.063426	0.064563	0.058491	0.061017
Gauss Mix Indep	0.061157	0.088023	0.063900	0.073316	0.057007	0.063081
NIG_factor	0.060878	0.078959	0.065270	0.070919	0.062097	0.062848

Table 2: ES 99%

References

Barbi, M. and S. Romagnoli. A copula-based quantile risk measure approach to estimate the optimal hedge ratio. *Journal of Futures Markets*, 34(7):658–675, 2014.

Barbi, M. and S. Romagnoli. A copula-based quantile risk measure approach to estimate the optimal hedge ratio. *Journal of Futures Markets*, 34(7):658–675, 2014.

Brandtner, M. and W. Kürsten. Decision making with expected shortfall and spectral risk measures: The problem of comparative risk aversion. *Journal of Banking & Finance*, 58:268–280, 2015.

	ERM k=10	ES 99%	ES 95%	VaR 99%	VaR 95%	Variance
Gaussian	0.034488	0.035237	0.034548	0.035123	0.034838	0.034248
t_Copula	0.034777	0.037100	0.035234	0.035634	0.035055	0.034494
t_Copula_Capped	0.034647	0.035679	0.034862	0.035282	0.034937	0.034322
Clayton	0.033714	0.037282	0.034230	0.036089	0.033445	0.034046
Frank	0.053661	0.047849	0.056299	0.053409	0.037638	0.046953
Gumbel	0.034028	0.035965	0.034528	0.036353	0.033568	0.034293
Plackett	0.034592	0.036831	0.035316	0.035752	0.034186	0.034558
Gauss Mix Indep	0.034439	0.045160	0.035120	0.040027	0.033756	0.034478
NIG_factor	0.033882	0.041001	0.035677	0.037975	0.034656	0.034453

Table 3: ES 95%

	ERM k=10	ES 99%	ES 95%	VaR 99%	VaR 95%	Variance
Gaussian	0.041327	0.044416	0.041943	0.043399	0.042275	0.041981
t_Copula	0.041450	0.044830	0.042806	0.043789	0.041693	0.041969
t_Copula_Capped	0.041498	0.044169	0.042411	0.044051	0.042018	0.042056
Clayton	0.040022	0.044523	0.042878	0.044215	0.040913	0.041943
Frank	0.076644	0.055387	0.081273	0.073433	0.046177	0.061056
Gumbel	0.042079	0.042139	0.042187	0.045340	0.040523	0.041937
Plackett	0.041013	0.044971	0.042370	0.042995	0.041574	0.041731
Gauss Mix Indep	0.040998	0.048017	0.043249	0.044518	0.040749	0.043386
NIG_factor	0.040457	0.047201	0.043925	0.044230	0.043240	0.043138

Table 4: VaR 99%

	ERM k=10	ES 99%	ES 95%	VaR 99%	VaR 95%	Variance
Gaussian	0.020385	0.020315	0.020143	0.020412	0.020121	0.019579
t_Copula	0.020547	0.020428	0.020661	0.020611	0.020370	0.019820
t_Copula_Capped	0.020525	0.020544	0.020503	0.020486	0.020224	0.019656
Clayton	0.019702	0.021042	0.020143	0.020640	0.019990	0.019700
Frank	0.026372	0.023529	0.027105	0.026212	0.023389	0.023594
Gumbel	0.019781	0.021311	0.020716	0.020421	0.019077	0.019541
Plackett	0.020459	0.020257	0.020589	0.020100	0.020237	0.020047
Gauss Mix Indep	0.020482	0.024753	0.020304	0.024158	0.019944	0.020723
NIG_factor	0.019923	0.023784	0.021009	0.022172	0.019980	0.020670

Table 5: VaR 95%

	ERM k=10	ES 99%	ES 95%	VaR 99%	VaR 95%	Variance
Gaussian	0.014387	0.014380	0.014360	0.014530	0.014670	0.014294
t_Copula	0.014378	0.014626	0.014343	0.014385	0.014627	0.014306
t_Copula_Capped	0.014375	0.014418	0.014332	0.014430	0.014643	0.014290
Clayton	0.014306	0.014870	0.014332	0.014532	0.014493	0.014267
Frank	0.021495	0.018982	0.022736	0.021476	0.018142	0.018897
Gumbel	0.014618	0.014971	0.014878	0.015438	0.014622	0.014321
Plackett	0.014444	0.014560	0.014424	0.014423	0.014596	0.014353
Gauss Mix Indep	0.014404	0.017404	0.014341	0.015671	0.014453	0.014408
NIG_factor	0.014362	0.015841	0.014484	0.015043	0.014474	0.014415

Table 6: Standard Deviation

Bühlmann, H. An economic premium principle. *ASTIN Bulletin: The Journal of the IAA*, 11(1):52–60, 1980.

Cecchetti, S. G., R. E. Cumby, and S. Figlewski. Estimation of the optimal futures hedge. *The Review of Economics and Statistics*, pages 623–630, 1988.

Creal, D., S. J. Koopman, and A. Lucas. A general framework for observation driven time-varying parameter models. Technical report, Tinbergen Institute Discussion paper, 2008.

Dowd, K., J. Cotter, and G. Sorwar. Spectral risk measures: properties and limitations. *Journal of Financial Services Research*, 34(1):61–75, 2008.

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

Engle, R. Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, 20(3):339–350, 2002.

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

Giacomini, E., W. Härdle, and V. Spokoiny. Inhomogeneous dependence modeling with time-varying copulae. *Journal of Business & Economic Statistics*, 27(2):224–234, 2009.

Hafner, C. M. and H. Manner. Dynamic stochastic copula models: Estimation, inference and applications. *Journal of Applied Econometrics*, 27(2):269–295, 2012.

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

Jürgen, F., W. Härdle, and C. Hafner. *Statistics of Financial Markets: An Introduction*. Springer, 2011.

Manner, H. and O. Reznikova. A survey on time-varying copulas: Specification, simulations, and application. *Econometric reviews*, 31(6):654–687, 2012.

Markowitz, H. Mean–variance approximations to expected utility. *European Journal of Operational Research*, 234(2):346–355, 2014.

McNeil, A., R. Frey, and P. Embrechts. *Quantitative Risk Management*. Princeton University Press, Princeton, NJ, 2005.

Nakamoto, S. Bitcoin: A peer-to-peer electronic cash system. Technical report, Manubot, 2019.

Patton, A. J. Modelling asymmetric exchange rate dependence. *International economic review*, 47(2):527–556, 2006.

6 Appendix

Proposition 5 *Let* $\boldsymbol{X} = (X_1, ..., X_d)^\top$ *be real-valued random variables with corresponding copula density* $\boldsymbol{c}_{X_1,...,X_d}$*, and continuous marginals* $F_{X_1}, ..., F_{X_d}$*. Then, density of the linear combination of marginals* $Z = n_1 \cdot X_1 + ... + n_d \cdot X_d$ *is*

$$f_Z(z) = \left|n_1^{-1}\right| \int_{[0,1]^{d-1}} \left[\boldsymbol{c}_{X_1,...,X_d}\{F_{X_1} \circ S(z), u_2, ..., u_d\} \cdot f_{X_1} \circ S(z)\right] du_2 ... du_d \tag{61}$$

$$S(z) = \frac{1}{n_1} \cdot z - \frac{n_2}{n_1} \cdot F_{X_2}^{-1}(u_2) - ... - \frac{n_d}{n_1} \cdot F_{X_d}^{-1}(u_d) \tag{62}$$

Proof. Rewrite $Z = n_1 \cdot X_1 + ... + n_d \cdot X_d$ in matrix

$$\begin{bmatrix} Y \\ X_2 \\ \vdots \\ X_d \end{bmatrix} = \begin{bmatrix} n_1 & n_2 & \cdots & n_d \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{bmatrix}. \tag{63}$$

By transformation of variable

$$\boldsymbol{f}_{Z,X_2,...,X_d}(z, x_2, ..., x_d) = \boldsymbol{f}_{X_1,...,X_d}\left(\boldsymbol{A}^{-1} \begin{bmatrix} z \\ x_2 \\ \vdots \\ x_d \end{bmatrix}\right) \cdot \left|\det \boldsymbol{A}^{-1}\right| \tag{64}$$

$$= \left|n_1^{-1}\right| \boldsymbol{f}_{X_1,...,X_d}\{S(z), x_2, ..., x_d\} \tag{65}$$

Let $u_i = F_{X_i}(x_i)$ and use the relationship

$$\boldsymbol{c}_{X_1,...,X_d}(u_1, ..., u_d) = \frac{\boldsymbol{f}_{X_1,...,X_d}(x_1, ..., x_d)}{\prod_{i=1}^d f_{X_i}(x_i)}, \tag{66}$$

we have

$$\boldsymbol{f}_{Z,X_2,...,X_d}(z, x_2, ..., x_d) = \tag{67}$$

$$\left|n_1^{-1}\right| \cdot \boldsymbol{c}_{X_1,...,X_d}\{F_{X_1} \circ S(z), u_2, ..., u_d\} \cdot f_{X_1}\{S(z)\} \cdot \prod_{i=2}^d f_{X_i}(x_i) \tag{68}$$

Result is obtained by integrating out $x_2, ...x_d$ by substitution $dx_i = \frac{1}{f_{X_i}(x_i)} du_i$. ■

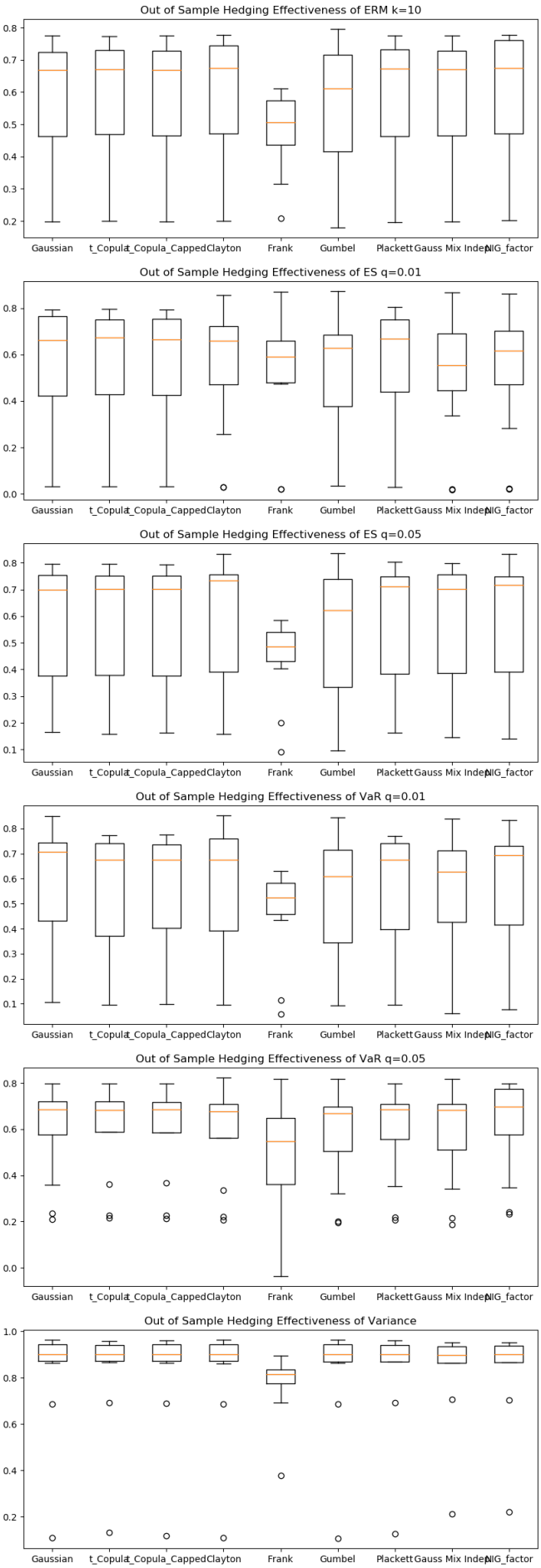

Figure 8: Out of Sample Hedging Effectiveness Box-plot. The HEs are obtained from a set of out-of-sample data, each set consists 30 days log returns of Bitcoin and CME future.

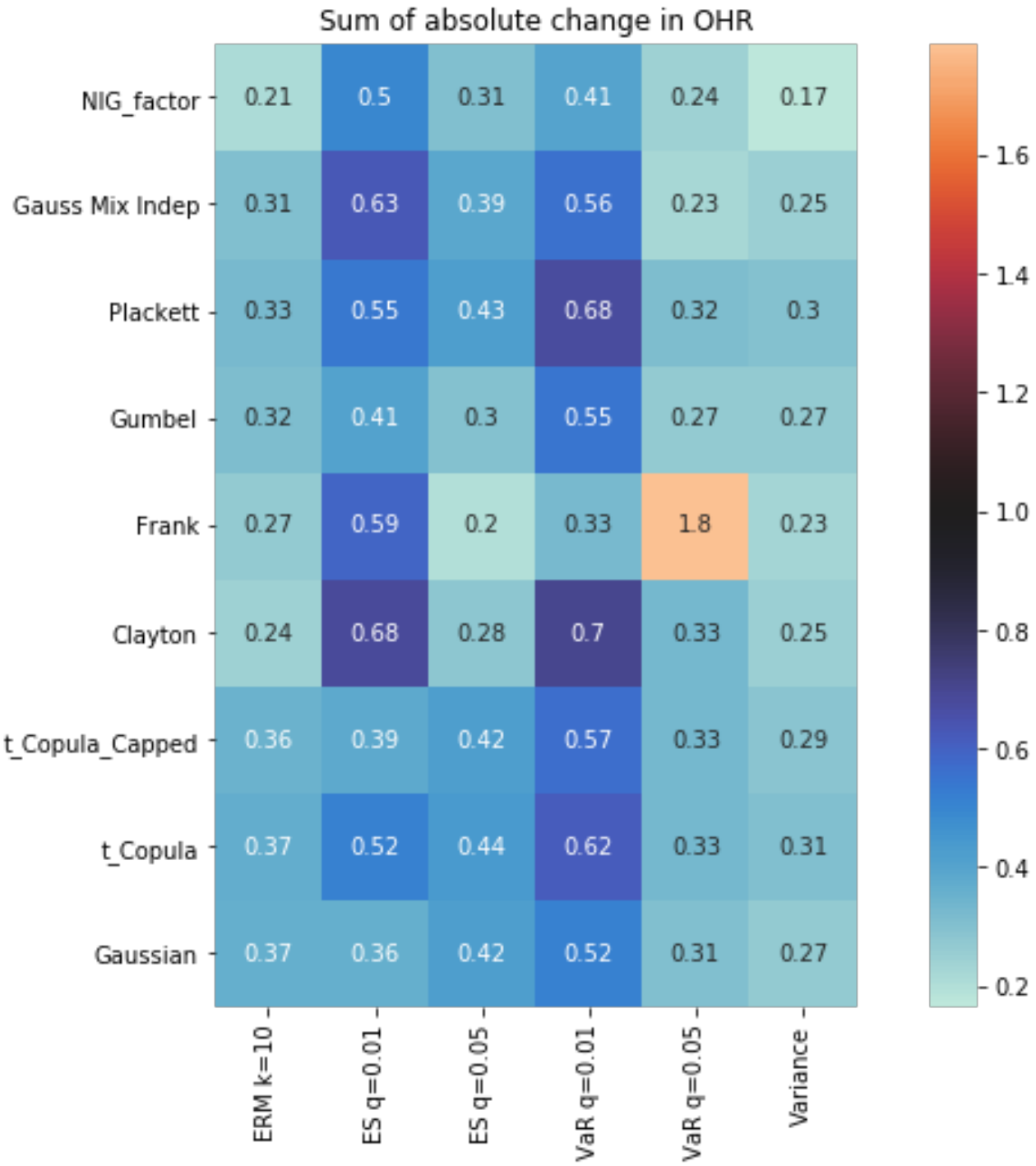

Figure 9: Sum of Absolute Change in OHR.