Hedging Futures with Spectral Risk Measures

This version: February 26, 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 ?. 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 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 ? 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 ?.

SRM is a weighted average of the quantiles 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 ?. 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 (?). However, it still causes some problems to capture the behaviors of investors when the value of absolute risk aversion (ARA) parameter beyond a threshold (?). If the relative risk aversion coefficients (RRA) are less than 1, ? 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 ? 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 (?). 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 quantiles 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. ? 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, ? 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 ?, 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 (?), 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: $h^* = \operatorname{argmin}_h \rho_\phi(s,h)$, 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).

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 (?)). 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*

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

Here $D_1C(u,v) = \dfrac{\partial}{\partial u}C(u,v)$, which is easily shown to fulfil, see e.g. Equation (5.15) of (**?**):[1]

$$D_1 C_{X,Y}(F_X(x), F_Y(y)) = \mathbf{P}(Y \le y | X = x). \tag{2}$$

Proof. Using the identity (2) 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 **?** we propose a more handy 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*

$$f_{r^h}(y) = \left|\frac{1}{h}\right| \int_0^1 c_{r^S, r^F}\left[u, F_{r^F}\left\{\frac{F_{r^S}^{-1}(u) - y}{h}\right\}\right] \cdot f_{r^F}\left\{\frac{F_{r^S}^{-1}(u) - y}{h}\right\} du \tag{3}$$

, or

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

The two expression 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 we have the copula density and the marginal densities. A generic expression can be found in the appendix (TODO).

2.1 Spectral Risk Measures

Spectral Risk Measures takes a form of

$$M_w(X) = -\int_0^1 w(p) q_s(X) ds \tag{5}$$

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: $w(p) \ge 0$.
- Increasingness: $w'(p) \ge 0$.

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

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

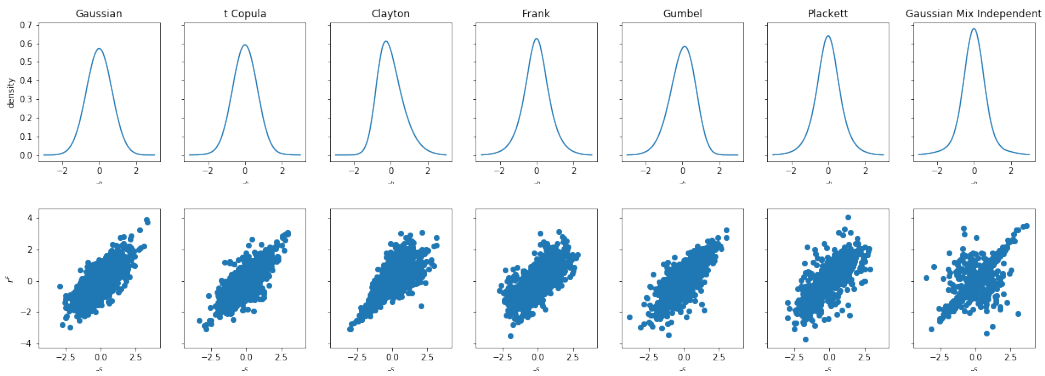

Figure 1: Upper Panel: Density of $Z = X - hY$ of different copulas with $X, Y \sim N(0,1)$, 0.75 Spearman's rho between X and Y, and $h = 0.5$; Lower Panel: Scatter plot of samples from copulas. This illustration shows how dependence structure modelled by copulas affects the density of the linear combination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

- 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 (?)

$$q_\alpha(X) = F_X^{-1}(\alpha), \quad \alpha \in (0,1). \tag{6}$$

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 (?). Expected shortfall (ES) measures are expressed as

$$\mathrm{ES}_\alpha(X) = \frac{1}{1-\alpha} \int_\alpha^1 q_s(X)ds \tag{7}$$

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 ?.

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:

$$w(p) = \frac{ke^{-k(1-p)}}{1-e^{-k}}, \quad \text{and} \quad W(p) = -\frac{1-e^{-k(1-p)}}{1-e^{-k}} \tag{8}$$

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 (5), the exponential SRM is,

$$M_w(X) = \int_0^1 \frac{ke^{-k(1-p)}}{1-e^{-k}} F^{-1}(p)dp \tag{9}$$

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,

$$w(p) = \gamma(1-p)^{\gamma-1}, \quad \text{and} \quad W(p) = -(1-p)^{\gamma} \tag{10}$$

Plugging the weighting function to (5), the power SRM is obtained,

$$M_w(X) = \int_0^1 \gamma(1-p)^{\gamma-1} F^{-1}(p)dp \tag{11}$$

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,

$$w(p) = \gamma p^{\gamma-1} \tag{12}$$

The Power SRM then becomes,

$$M_w(X) = \int_0^1 \gamma p^{\gamma-1} F^{-1}(p)dp \tag{13}$$

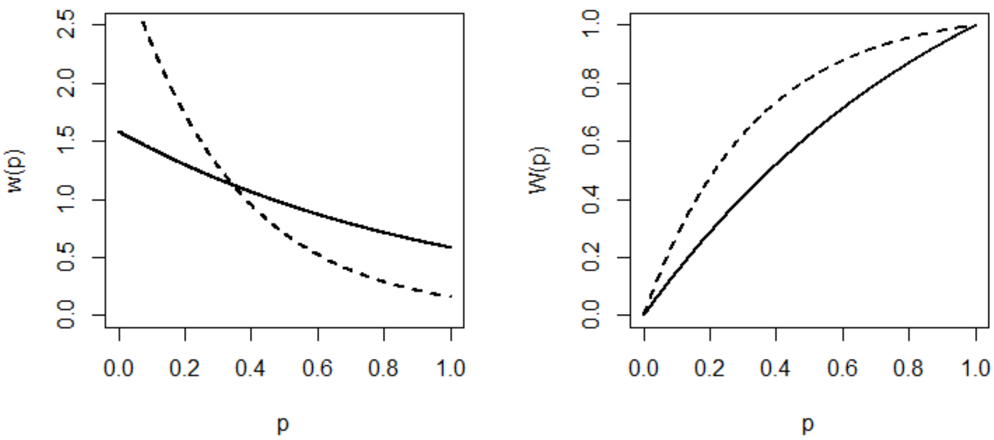

Figure 2: Exponential SRMs for $k = 1$ (dashed) and $k = 2$ (solid).

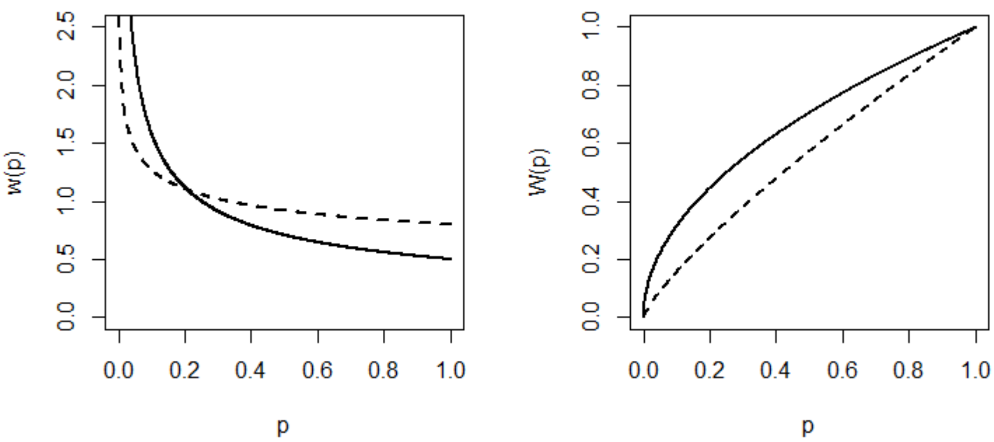

Figure 3: Power SRMs for $\gamma = 0.5$ (solid) and $\gamma = 0.8$ (dashed).

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.

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 (B1 in Joe) has a form

$$\boldsymbol{C}(u,v) = \Phi_{2,\rho}\{\Phi^{-1}(u), \Phi^{-1}(v)\} \tag{14}$$

$$= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right) dsdt \tag{15}$$

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 copula density of Gaussian copula can be written as

$$\boldsymbol{c}_\rho(u,v) = \frac{\phi_{2,\rho}\{\Phi^{-1}(u), \Phi^{-1}(v)\}}{\phi\{\Phi^{-1}(u)\} \cdot \phi\{\Phi^{-1}(v)\}} \tag{16}$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{u^2 - 2\rho uv + v^2}{2(1-\rho^2)}\right), \tag{17}$$

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

$$\tau_K(\rho) = \frac{2}{\pi} \arcsin \rho \tag{18}$$

$$\rho_S(\rho) = \frac{6}{\pi} \arcsin \frac{\rho}{2} \tag{19}$$

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

$$\boldsymbol{C}(u,v) = \boldsymbol{T}_{2,\rho,\nu}\{T_\nu^{-1}(u), T_\nu^{-1}(v)\} \tag{20}$$

$$= \int_{-\infty}^{T_\nu^{-1}(u)} \int_{-\infty}^{T_\nu^{-1}(v)} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\pi\nu\sqrt{1-\rho^2}} \tag{21}$$

$$\left(1 + \frac{s^2 - 2st\rho + t^2}{\nu}\right)^{-\frac{\nu+2}{2}} dsdt, \tag{22}$$

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

$$\boldsymbol{c}(u,v) = \frac{\boldsymbol{t}_{2,\rho,\nu}\{T_\nu^{-1}(u), T_\nu^{-1}(v)\}}{t_\nu\{T_\nu^{-1}(u)\} \cdot t_\nu\{T_\nu^{-1}(v)\}}, \tag{23}$$

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

$$\boldsymbol{C}(u,v) = \psi^{-1}\{\psi(u), \psi(v)\}, \tag{24}$$

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.

This section will briefly introduce this class of copula, we refer readers to Nelsen and Joe for detail of this class of copula.

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

$$\boldsymbol{C}_\theta(u,v) = \frac{1}{\log(\theta)} \log\left\{1 + \frac{(\theta^u - 1)(\theta^v - 1)}{\theta - 1}\right\} \tag{25}$$

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

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

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

$$\tau_K(\theta) = 1 - 4\frac{D_1\{-\log(\theta)\}}{\log(\theta)}, \tag{27}$$

and

$$\rho_S(\theta) = 1 - 12\frac{D_2\{-\log(\theta)\} - D_1\{\log(\theta)\}}{\log(\theta)}, \tag{28}$$

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.

$$\boldsymbol{C}_\theta(u,v) = \exp -\{(-\log(u))^\theta + (-\log(v))^\theta\}^{\frac{1}{\theta}}, \tag{29}$$

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.

$$\tau_K(\theta) = \frac{\theta - 1}{\theta} \tag{30}$$

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

$$\boldsymbol{C}_\theta(u, v) = \Big[\max\{u^{-\theta} + v^{-\theta} - 1, 0\}\Big]^{-\frac{1}{\theta}}, \tag{31}$$

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

$$\tau_K(\theta) = \frac{\theta}{\theta + 2}. \tag{32}$$

Plackett copula has an expression

$$\boldsymbol{C}_\theta(u, v) = \frac{1 + (\theta - 1)(u + v)}{2(\theta - 1)} - \frac{\sqrt{\{1 + (\theta - 1)(u + v)\}^2 - 4uv\theta(\theta - 1)}}{2(\theta - 1)} \tag{33}$$

$$\rho_S(\theta) = \frac{\theta + 1}{\theta - 1} - \frac{2\theta \log \theta}{(\theta - 1)^2} \tag{34}$$

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

$$\begin{aligned}
&\frac{\mathbb{P}(U \le u, V \le v) \cdot \mathbb{P}(U > u, V > v)}{\mathbb{P}(U \le u, V > v) \cdot \mathbb{P}(U > u, V \le v)} && (35)\\
&= \frac{C_\theta(u, v)\{1 - u - v + C_\theta(u, v)\}}{\{u - C_\theta(u, v)\}\{v - C_\theta(u, v)} && (36)\\
&= \theta. && (37)
\end{aligned}$$

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.

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

$$\mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) = \sum_{k=1}^{K} p^k \cdot \boldsymbol{C}^{(k)}\{F_{X_1}^{(k)}(x_1; \boldsymbol{\gamma}_1^{(k)}), F_{X_2}^{(k)}(x_2; \boldsymbol{\gamma}_2^{(k)}); \boldsymbol{\theta^{(k)}}\} \tag{38}$$

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

$$\boldsymbol{C}(u, v) = \sum_{k=1}^{K} p^{(k)} \cdot \boldsymbol{C}^{(k)}\{F_{X_1}^{-1}(u), F_{X_2}^{-1}(v); \boldsymbol{\theta^{(k)}}\}. \tag{39}$$

The copula density is again a linear combination of copula density

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

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*

$$\rho_S = \sum_{k=1}^{K} p^{(k)} \cdot \rho_S^{(k)} \tag{41}$$

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

$$\rho_S = 12 \int_{\mathbb{I}^2} \boldsymbol{C}(s, t) ds dt - 3. \tag{42}$$

Rewrite the mixture copula into sumation of components

$$\rho_S = 12 \int_{\mathbb{I}^2} \sum_{k=1}^{K} p^{(k)} \cdot \boldsymbol{C}^{(k)}(s, t) ds dt - 3. \tag{43}$$

■

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)*

$$\boldsymbol{C}_{\alpha,\beta}(u, v) = \alpha \boldsymbol{M}(u, v) + (1 - \alpha - \beta)\boldsymbol{\Pi}(u, v) + \beta \boldsymbol{W}(u, v), \tag{44}$$

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

$$\tau_K(\alpha, \beta) = \frac{(\alpha - \beta)(\alpha + \beta + 2)}{3} \tag{45}$$

, and

$$\rho_S(\alpha, \beta) = \alpha - \beta \tag{46}$$

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

$$\boldsymbol{C}(u, v) = p \cdot \boldsymbol{C}^{\text{Gaussian}}(u, v) + (1 - p)(uv). \tag{47}$$

The corresponding copula density is

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

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.

4 Estimation

4.1 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

$$\rho_S = 12 \int\int_{I^2} C_{\boldsymbol{\theta}}(u, v)\, \mathrm{d}u\, \mathrm{d}v - 3 \tag{49}$$

$$\tau_K = 4\mathbb{E}[C_{\boldsymbol{\theta}}\{F_X(x), F_Y(y)\}] - 1, \tag{50}$$

$$\lambda_q = \begin{cases} \mathbf{P}(F_X(X) \leq q | F_Y(Y) \leq q) = \dfrac{C_{\boldsymbol{\theta}}(q, q)}{q}, & \text{if } q \in (0, 0.5], \\ \mathbf{P}(F_X(X) > q | F_Y(Y) > q) = \dfrac{1 - 2q + C_{\boldsymbol{\theta}}(q, q)}{1 - q}, & \text{if } q \in (0.5, 1). \end{cases} \tag{51}$$

The empirical counterparts are

$$
\begin{aligned}
\hat{\rho}_S &= \frac{12}{n}\sum_{k=1}^{n}\hat{F}_X(x_k)\hat{F}_Y(y_k) - 3,\\
\hat{\tau}_K &= \frac{4}{n}\sum_{k=1}^{n}\hat{C}\{\hat{F}_X(x_i),\hat{F}_X(y_i)\} - 1,\\
\hat{\lambda}_q &= \begin{cases} \dfrac{1}{n}\displaystyle\sum_{k=1}^{n}\frac{\mathbf{1}_{\{\hat{F}_X(x_k)\le q,\hat{F}_Y(y_k)\le q\}}}{q}, & \text{if } q\in(0,0.5],\\ \dfrac{1}{n}\displaystyle\sum_{k=1}^{n}\frac{\mathbf{1}_{\{\hat{F}_X(x_k)> q,\hat{F}_Y(y_k)> q\}}}{1-q}, & \text{if } q\in(0.5,1). \end{cases},
\end{aligned}
$$

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

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.

4.2 Maximum Likelihood Estimation

The MLE estimator for a bivariate parametric copula is

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmax}}\, l(X,Y;\boldsymbol{\theta}), \tag{52}$$

where

$$l(X,Y;\boldsymbol{\theta}) = \sum_{k=1}^{n}\log c(x_i,y_i;\boldsymbol{\theta}). \tag{53}$$

Procedure of maximising equation 52 as a whole is called exact maximum likelihood method. Leveraging the attractive feature of copula that one can model the dependence structure and marginals separately, we rewrite 53 into canonical expression

$$l(X,Y;\boldsymbol{\theta}) = \sum_{k=1}^{n}\log c\{F_X(x_i;\delta_X),F_Y(y_i;\delta_Y);\boldsymbol{\gamma}\} + \sum_{k=1}^{n}\log f_X(x_i;\boldsymbol{\delta}_X) + \sum_{k=1}^{n}\log f_X(y_i;\boldsymbol{\delta}_Y), \tag{54}$$

where the $\boldsymbol{\gamma}$ is the dependence parameter in the copula and $\boldsymbol{\delta}$ is the parameters in the margins.

The inference-functions for margins (IFM) approach by Joe is a two step procedure of maximising 52. The approach calibrate first the $\boldsymbol{\delta}$s and then the $\boldsymbol{\gamma}$.

Similar to the IFM approach, pseudo-maximum likelihood approach by Genest and Rivest (1993) replace the parametric margins by empirical estimates, we rewrite 53 again with

$$l(X,Y;\boldsymbol{\theta}) = \sum_{k=1}^{n} \log c(u_i, v_i; \boldsymbol{\gamma}), \tag{55}$$

where $u_i = \hat{F}_X(x_i)$ and $v_i = \hat{F}_Y(y_i)$.

4.3 Comparison

Both the simulated method of moments and the maximum likelihood estimation are unbiased and proven to give good fits (ref ...). 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 quantile dependence.

5 Data and margins

TODO:

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

6 Results

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

6.1 Hedging Effectiveness

The hedging effectiveness(HE) is defined as

$$1 - \frac{\rho_\phi(R^h)}{\rho_\phi(R^S)}. \tag{56}$$

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 ... 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

$$1 - \frac{\mathrm{ES}_\alpha(R^h)}{\mathrm{ES}_\alpha(R^S)}. \tag{57}$$

Figure 5 shows the out-of-sample hedging effectiveness of difference copulas under various risk reduction objectives. Observe that in every 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 every copula. In some instances, the HE can be as low as 10%.

6.2 Ability of hedging extreme negative events

Figure 6 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.

6.3 Stability of h^*

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

$$\sum_{t=1}^{T} |h_t - h_{t-1}|. \tag{58}$$

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.

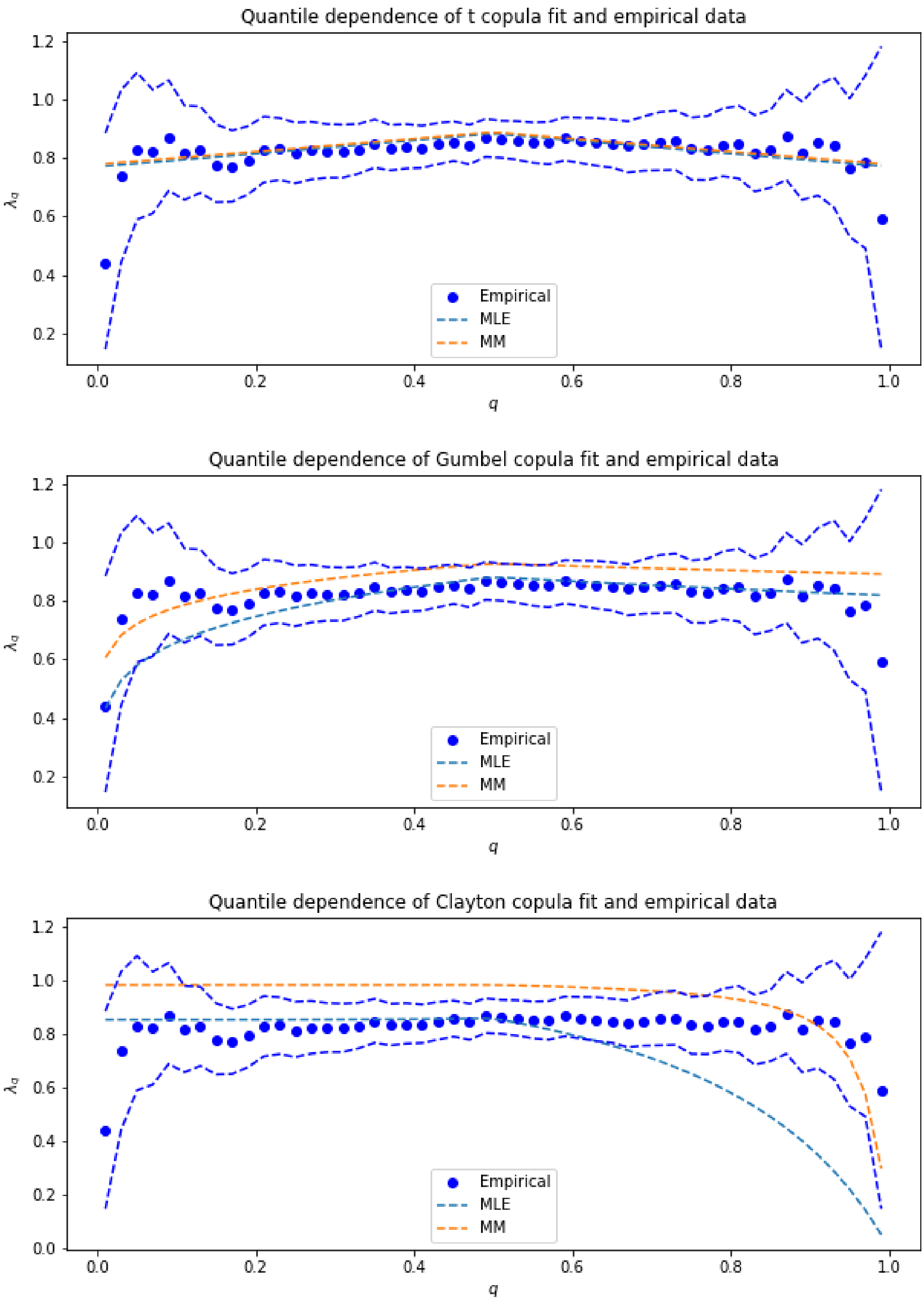
Quantile dependence of t copula fit and empirical data
Empirical
MLE
MM
q
Quantile dependence of Gumbel copula fit and empirical data
Empirical
MLE
MM
q
Quantile dependence of Clayton copula fit and empirical data
Empirical
MLE
MM
q

Figure 4

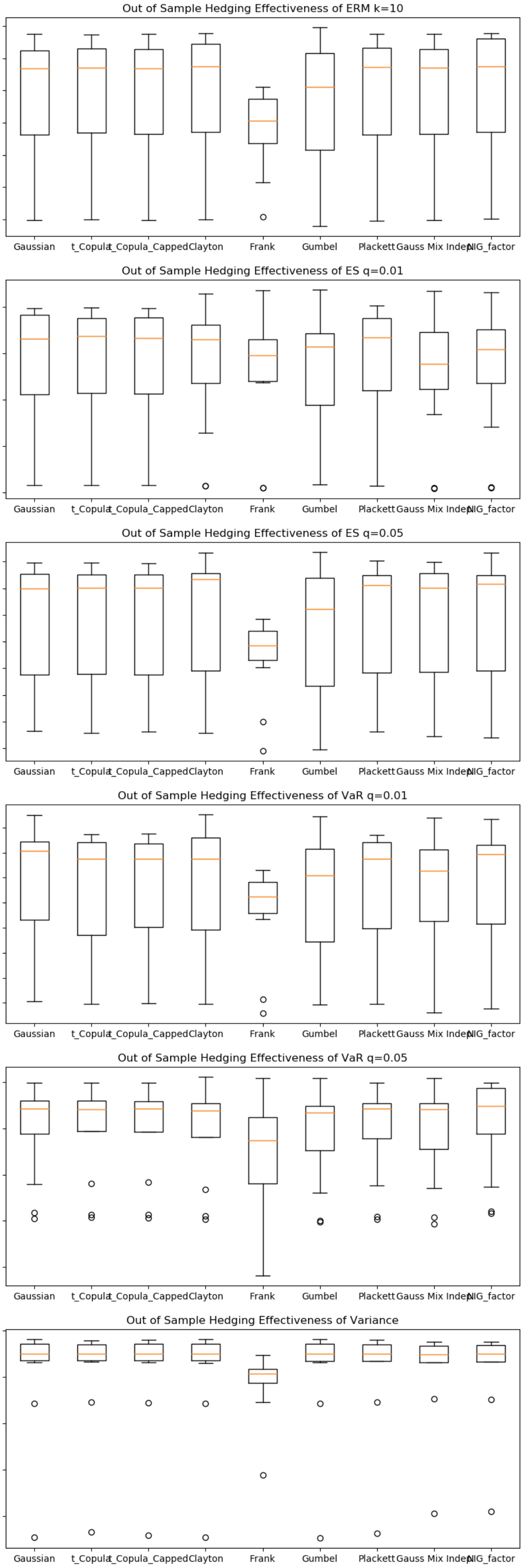

Figure 5: Out of Sample Hedging Effectiveness Boxplot. The HEs are obtained from a set of out-of-sample data, each set consists 30 days log returns of Bitcoin and CME future.

Figure 6: Upper Panel: Out of Sample Log-return of Bitcoin; Lower Panel: Out of Sample Hedged Portfolio log-returns. The h^* is obtained from Gumbel copula aiming at reducing variance. The red dots indicate the 30 most extreme negative returns in Bitcoin.

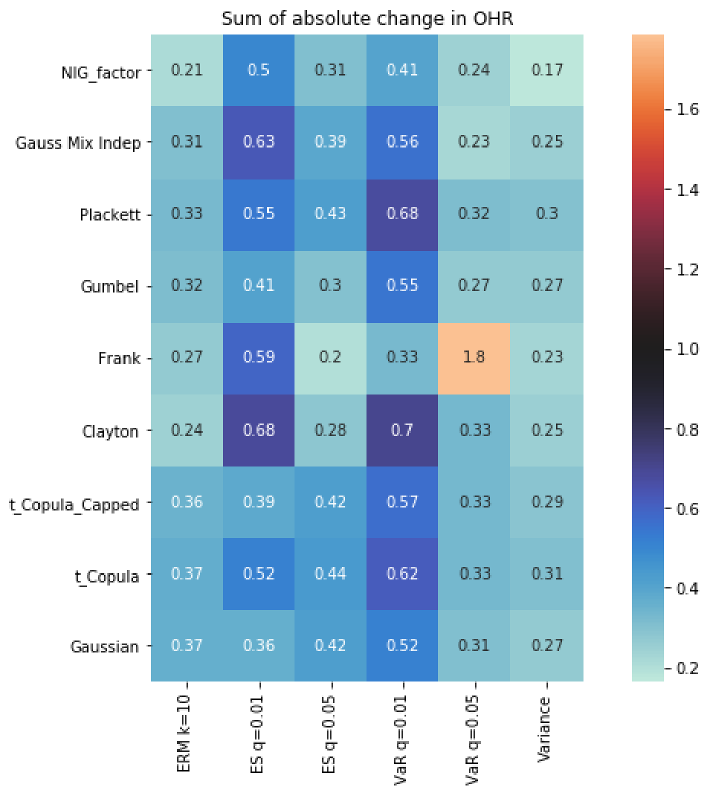

Figure 7: Sum of Absolute Change in OHR.