Hedging Cryptos with Bitcoin Futures

Francis Liu*

Meng-Jou Lu †

Natalie Packham‡

Wolfgang Karl Härdle§¶

This version: May 11, 2021

Abstract

The introduction of derivatives on Bitcoin, in particular the launch of futures contracts on CME in December 2017 and introduction of cryptocurrency indices like the CRIX or the Bloomberg Galaxy Crypto Index enables investors to hedge risk exposures of Bitcoin by futures or contingent claims on indices. We investigate methods of finding the optimal hedge ratio h^* under different dependence structures modeled by copulae and optimality definition described by a range of risk measures. 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 a variety of Spectral Risk Measures. In addition, we deploy different copulae for capturing the dependency between spot and future returns, such as the Gaussian, Student-t, NIG and Archimedean copulae. Various measures of hedge effectiveness in out-of-sample tests give insights in the practice of hedging Bitcoin. We find that across copulae and risk measures, the hedge effectiveness are very similar with the exception of the Frank copula, Expected Shortfall 99% and Value-at-Risk 99%. Our findings are based on an analysis for the time span from 15/12/2017 to 04/02/2021.

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

---

*Department of Business and Economics, Berlin School of Economics and Law, Badensche Str. 52, 10825 Berlin, Germany. Blockchain Research Center, Humboldt-Universität zu Berlin, Germany. International Research Training Group 1792, Humboldt-Universität zu Berlin, Germany. E-mail: Francis.Liu@hwr-berlin.de.

†Department of Finance, Asia University, 500, Lioufeng Rd., Wufeng, Taichung 41354, Taiwan Department of Finance, Asia University, 500, Lioufeng Rd., Wufeng, Taichung 41354, Taiwan E-mail: mangrou@gmail.com.

‡Department of Business and Economics, Berlin School of Economics and Law, Badensche Str. 52, 10825 Berlin, Germany. International Research Training Group 1792, Humboldt-Universität zu Berlin, Germany. E-mail: packham@hwr-berlin.de.

§Blockchain Research Center, Humboldt-Universität zu Berlin, Germany. Wang Yanan Institute for Studies in Economics, Xiamen University, China. Sim Kee Boon Institute for Financial Economics, Singapore Management University, Singapore. Faculty of Mathematics and Physics, Charles University, Czech Republic. National Yang Ming Chiao Tung University, Taiwan. E-mail: haerdle@wiwi.hu-berlin.de.

¶Financial support of the European Union's Horizon 2020 research and innovation program "FIN-TECH: A Financial supervision and Technology compliance training programme" under the grant agreement No 825215 (Topic: ICT-35-2018, Type of action: CSA), the European Cooperation in Science & Technology COST Action grant CA19130 - Fintech and Artificial Intelligence in Finance - Towards a transparent financial industry, the Deutsche Forschungsgemeinschaft's IRTG 1792 grant, the Yushan Scholar Program of Taiwan, the Czech Science Foundation's grant no. 19-28231X / CAS: XDA 23020303, as well as support by Ansar Aynetdinov (ansar.aynetdinov@hu-berlin.de) are greatly acknowledged.

1 Introduction

Cryptocurrencies (CCs) are a growing asset class. Many more CCs are now available on the market since the first cryptocurrency Bitcoin (BTC) surfaced, Nakemoto (2009). As the network effect weighs in, the prices of bitcoin and its variants have risen in tandem. These innovations and the perceived investment potential have led to rapid growth in the number of altcoins and the market size of CC. The price of bitcoin even surged to USD 20,000 at beginning of 2021. Bitcoin is popular with the techno tribe, the currency is regarded as being beyond the reach of government regulation- the anonymous founder of Bitcoin introduced the idea of a distributed block chain to prevent the counterfeiting of Bitcoin (Oet et al., 2015). In response to the rapid development of the CC market, the CME group launched a BTC future contract in Dec 2017. Trading volume in BTC futures surpassed $2 trillion in 2020 (CryptoCompare, 2020). While more and more investors (individuals and institution) are adding CCs and their derivatives into their portfolios. we see the need to understand the downside risk and find a suitable way to hedge and are interested in resisting extreme risks and improve their profits. This paper analyse modern techniques for the choice of the hedge ratio of the CC portfolios with various copulae and risk measures.

Optimal hedge ratio is the appropriate size of futures contracts which should be held such that the movements in future price cancel that of BTC. The task of determining an optimal hedge ratio is not easy. It relies on the dependence between the BTC price and future price. Copulae provide the flexibility to model multivariate random variable separately by its margins and dependence structure. The concept of copulae was originated by Wassily Hoeffding (Fisher and Sen, 2012) and, despite the slight difference, popularised by the work of Abe Sklar (Sklar, 1959). Different risk measures account for investors' risk attitude. They serve as loss functions in the searching process of optimal hedge ratio. Vast literature discussed the relationship between risk measures and investor's risk attitude, we refer readers to Artzner et al. (1999) for an axiomatic, economic reasoning approach of risk measure construction; Embrechts et al. (2002) for reasoning of using Expected Shortfall (ES) and Spectral Risk Measure (SRM) in addition to VaR; Acerbi (2002) for direct linkage between risk measures and investor's risk attitude using the concept of "risk aversion function".

Financial asset return is known to be non Gaussian (Fama, 1963). In particular, Gaussian models cannot produce so-called fat tails and asymmetry of observed probability densities, which leads to underestimate financial risks. Therefore, one cannot solely rely on 2^{nd} order moment calculations in order to minimize downside risk. Variance as a risk measure doesn't consider the variety of investors' utility functions. However, the investors are tail-risk averse. Bollerslev et al. (2015) find that the jump tail risk is more closely associated with changes in risk-aversion. It is important to link the investor utility's functions as hedging the tail risk. Significant tail risks lead to the need to investigate even static hedge with more refined methods than minimum-variance based (Ederington and Salas, 2008).

In order to capture the risk preferences of investors, in addition to variance, we include other risk measures. We consider also Value-at-Risk (VaR), Expected Shortfall (ES), and Spectral Risk Measure (SRM). VaR is widely used by the industry and easy to understand. ES and SRM are chosen because of their coherence property, in particular, they encourage diversification. SRM is also directly related to individual's utility function. Popular examples are the exponential SRM and power SRM introduced by Dowd et al. (2008).

This paper considers hedging Bitcoin using its future. i.e. to find an optimal hedge ratio h^* such that the risk of a hedged portfolio $r^h = r^S - h^* r^F$ has minimal risk. Here r^S as the log return of BTC spot price, r^F the log return of Bitcoin future. The leptokurtic properties mentioned above leads us to deploy a comprehensive way of modelling dependency namely copulae together with various risk measures as loss function to find optimal hedge ratio. In this paper, we first calibrate the log returns of BTC and CME future by copulae, then find the optimal quantity of assets in the hedged portfolio according to a range of risk measures. Barbi and Romagnoli (2014) use the C-convolution operator introduced by Cherubini et al. (2011) to derive the distribution of linear combination of margins with copula as their dependence structure. Our main result shows that the distribution function of a linear combination of random variables can be expressed via the copula and margins.

This research proposes the model techniques for the analysis of the hedging strategy on the CC's tail risk in five aspects. The remainder of the article is organized as follows. Section 2 methodology. Section 3 data, and Section 4 empirical result. Section 5 concludes.

All calculations in this work can be reproduced. The codes are available on www.quantlet.com.

2 Optimal hedge ratio

We form a portfolio with two assets, a spot asset and a future contract, for example Bitcoin spot and CME Bitcoin future. Our objective is to minimize the risk of the exposure in the spot. To keep a simple portfolio setting, we long one unit of the spot and short h unit of the future with $h \in [0, \infty)$. Let r^S and r^F be the log returns of the spot and future price, the log return of the portfolio is

$$r^h = r^S - hr^F.$$

We call this portfolio a hedged portfolio: the price movement of spot is hedged by the price movement of future.

Risk is measured by risk measures. Assume the payoff r^h of a portfolio lives in a probability space, $r^h \in L(\Omega, \mathcal{F}, \mathbb{P})$, and there is a risk measure on r^h $\rho : r^h \mapsto \mathbb{R}$. We are looking for an optimal hedge ratio h^* which minimizes risk measure

$$h^* = \underset{h}{\operatorname{argmin}}\, \rho(r^h).$$

Most risk measures are defined as functionals of the portfolio loss distribution F_{r^h}, i.e. $\rho : F_{r^h} \mapsto \mathbb{R}$. For example, Value-at-Risk (VaR) is simply the quantile of r^h multiply with negative one $\mathrm{VaR}_{1-\alpha} = -F_{r^h}^{(-1)}(1-\alpha) = -\inf\{x \in \mathbb{R} : 1-\alpha \leq F_{r^h}(x)\}$, where α is a parameter chosen by investor. We need the knowledge of F_{r^h} in order to measure risk. By convolution of random variables (Härdle and Léopold, 2011), $f_{r^h}(z) = \int_{-\infty}^{\infty} f_{r^S,-hr^F}(x, z-x)dx$, where $f_{r^S,-hr^F}$ is the joint pdf of r^S and $-hr^F$. Obviously the cdf of r^h and risk measure depend on the joint distribution of r^S and $-hr^F$.

Optimising h according to $f_{r^S,-hr^F}$ is unfavorable in a sense that one needs to calibrate a new joint pdf $f_{r^S,-hr^F}$ when updating h. This is too time consuming and unnecessary. Another problem of using joint pdf is that one lacks of flexibility to model the margins. A joint pdf completely determine the form of its marginals, for example, margins of a bivariate t-distribution are univariate t-distributions.

To overcome the problems, we use copulae. The benefit of using copulae is two folded. First, copulae allow us to model the margins and dependence structure separately, see Sklar's Theorem. Second, copulae are invariance under strictly monotone increasing function (Schweizer et al., 1981), see lemma below.

Theorem 1 (Sklar's Theorem) *Let F be a joint distribution function with margins F_X, F_Y. Then, there exists a copula $C : [0,1]^2 \mapsto [0,1]$ such that, for all $x, y \in \mathbb{R}$*

$$F(x, y) = C\{F_X(x), F_Y(y)\}. \tag{1}$$

If the margins are continuous, then C is unique; otherwise C is unique on the range of the margins.

Conversely, if C is a copula and F_X, F_Y are univariate distribution functions, then the function F defined by (1) is a joint distribution function with margins F_X, F_Y.

Indeed, many basic results about copulae can be traced back to early works of Wassily Hoeffding (Hoeffding, 1940, 1941). The works aimed to derive a measure of relationship of variables which is invariant under change of scale. Readers can refer to Fisher and Sen (2012) for English translations of the works.

Lemma 1

$$C_{X,hY}\{F_X(s), F_{hY}(t)\} = C_{X,Y}\{F_X(s), F_Y(t/h)\}. \tag{2}$$

Leveraging the two features of copulae, Barbi and Romagnoli (2014) introduces the distribution of linear combination of random variables using copulae. We slightly edit the Corollary 2.1 of their work and yield the following correct expression of the distribution.

Proposition 2 *Let X and Y be two real-valued continuous random variables on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with absolutely continuous copula $C_{X,Y}$ and marginal distribution functions F_X and F_Y. Then, the distribution function of of Z is given by*

$$F_Z(z) = 1 - \int_0^1 D_1 C_{X,Y}\left[u, F_Y\left\{\frac{F_X^{(-1)}(u) - z}{h}\right\}\right] du. \tag{3}$$

Here, $F^{(-1)}$ denotes the inverse of F, i.e., the quantile function.

Here $D_1 C(u, v) = \frac{\partial}{\partial u} C(u, v)$ see e.g. Equation (5.15) of (McNeil et al., 2005):

$$D_1 C_{X,Y}(F_X(x), F_Y(y)) = \mathbf{P}(Y \leq y | X = x). \tag{4}$$

Proof. Using the identity (4) gives

$$\begin{aligned}
F_Z(z) &= \mathbf{P}(X - hY \leq z) = \mathsf{E}\left\{\mathbf{P}\left(Y \geq \frac{X - z}{h}\Big| X\right)\right\} \\
&= 1 - \mathsf{E}\left\{\mathbf{P}\left(Y \leq \frac{X - z}{h}\Big| X\right)\right\} = 1 - \int_0^1 D_1 C_{X,Y}\left[u, F_Y\left\{\frac{F_X^{(-1)}(u) - z}{h}\right\}\right] du.
\end{aligned}$$

■

In addition to Barbi and Romagnoli (2014) we propose a more handy expression for the pdf of Z.

Corollary 1 *Given the formulation of the above portfolio, the pdf of Z can be written as*

$$f_Z(z) = \left|\frac{1}{h}\right| \int_0^1 c_{X,Y}\left[F_Y\left\{\frac{F_X^{(-1)}(u) - z}{h}\right\}, u\right] \cdot f_Y\left\{\frac{F_X^{(-1)}(u) - z}{h}\right\} du \tag{5}$$

, *or*

$$f_Z(z) = \int_0^1 c_{X,Y} \left[F_X \left\{ z + hF_Y^{(-1)}(u) \right\}, u \right] \cdot f_X \left\{ z + hF_Y^{(-1)}(u) \right\} du. \tag{6}$$

The two expression are equivalent. Notice that the pdf of Z in the above proposition is readily accessible as long as we have the copula density and the marginal densities. The proof and a generic expression can be found in the appendix.

2.1 Risk Measures

We consider four risk measures: variance, Value-at-Risk (VaR), Expected Shortfall (ES), and Exponential Risk Measure (ERM). A summary of risk measures being used in portfolio selection problem can be found in Härdle et al. (2008).

Let Z be a random variable of distribution F_Z.

- Variance is $\mathrm{Var}(F_Z) = \int_{\mathbb{R}} z^2 dF_Z(z)$
- VaR of a given confidence level α is $\mathrm{VaR}(F_Z) = -F_Z^{(-1)}(1-\alpha)$
- ES with parameter α is $\mathrm{ES}(F_Z) = -\frac{1}{1-\alpha} \int_0^{1-\alpha} F_Z^{(-1)}(p)dp$
- ERM with Arrow-Pratt coeficient of absolute risk aversion k is $\mathrm{ERM}_k(F_Z) = \int_0^{1-\alpha} \phi(p) F_Z^{(-1)}(p)dp$ where ϕ is a weight function describe in equation 8 below.

VaR, ES, and ERM fall into the class of Spectral Risk Measure (SRM). SRM has the form (Acerbi, 2002)

$$\rho_\phi(r^h) = -\int_0^1 \phi(p) F_Z^{(-1)}(p)dp, \tag{7}$$

where p is the loss quantile and $\phi(p)$ is a user-defined weighting function defined over $[0,1]$. We consider only admissible risk spectrum *[named by Acerbi (2002)]* $\phi(p)$:

- ϕ is positive
- ϕ is decreasing
- $\int_{[0,1]} \phi(q)dq = 1$

The VaR's $\phi(p)$ gives all its weight on the $1-\alpha$ quantile of Z and zero elsewhere, i.e. the weighting function is a Dirac delta function. The ES' $\phi(p)$ gives all tail quantiles the same weight of $\frac{1}{1-\alpha}$ and non-tail quantiles zero weight. ERM assumes investor's risk preference is in a form of exponential utility function $U(x) = -e^{kx}$, its risk spectrum is defined as

$$\phi(p) = \frac{ke^{-k(1-p)}}{1-e^{-k}}, \tag{8}$$

where k is the Arrow-Pratt coefficient of absolute risk aversion.

k has an economic interpretation of being the ratio between the second derivative and first derivative of investor's utility function on an risky asset

$$k = -\frac{U''(x)}{U'(x)}, \tag{9}$$

for x in all possible outcomes.

2.2 Copulae

As we saw from the last section, risk measures we considered are all functionals of joint distribution of r^S and r^F. We test different copulae for their ability to model joint distribution of crypto-related assets returns. We consider Gaussian-, t-, Frank-, Gumbel-, Clayton-, Plackett-, mixture, and factor copula. This hedging exercise concerns only portfolios with two assets, we only present the bivariate version of copulae and some important features of a copula, they include Kendall's τ, Spearman's ρ, upper tail dependence $\lambda_U \stackrel{\text{def}}{=} \lim_{q \to 1^-} \mathbf{P}\{X > F_X^{(-1)}(q) | Y > F_Y^{(-1)}(q)\}$ and lower tail dependence $\lambda_L \stackrel{\text{def}}{=} \lim_{q \to 0^+} \mathbf{P}\{X \leq F_X^{(-1)}(q) | Y \leq F_Y^{(-1)}(q)\}$. Furthermore, we denote the Fréchet-Hoeffding lower bound as $\boldsymbol{W}$, product copula as $\boldsymbol{\Pi}$, and the Fréchet-Hoeffding upper bound as $\boldsymbol{M}$, they represent cases of perfect negative dependence, independence, and perfect positive dependence respectively. For further detail, we refer readers to Joe (1997) and Nelsen (1999). See also Härdle and Okhrin (2010).

2.2.1 Elliptical Copulae

Elliptical copulae are dependence structure associated with elliptical distributions. The bivariate Gaussian copula is:

$$\begin{aligned}\boldsymbol{C}(u,v) &= \Phi_{2,\rho}\{\Phi^{-1}(u), \Phi^{-1}(v)\} \\ &= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right) ds dt \end{aligned} \tag{10}$$

where $\Phi_{2,\rho}$ is the cdf of bivariate Normal distribution with zero mean, unit variance, and correlation ρ *the rho we use in risk measure is $\rho_{something}$. This should remove ambiguity.*, and Φ^{-1} is quantile function univariate standard normal distribution. The Gaussian copula density is

$$\begin{aligned}\boldsymbol{c}_\rho(u,v) &= \frac{\boldsymbol{\varphi}_{2,\rho}\{\Phi^{-1}(u), \Phi^{-1}(v)\}}{\varphi\{\Phi^{-1}(u)\} \cdot \varphi\{\Phi^{-1}(v)\}} \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{u^2 - 2\rho uv + v^2}{2(1-\rho^2)}\right), \end{aligned} \tag{11}$$

where $\boldsymbol{\varphi}_{2,\rho}(\cdot)$ is the pdf of $\Phi_{2,\rho}$, and $\varphi(\cdot)$ the standard normal distribution pdf.

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

$$\tau_K(\rho) = \frac{2}{\pi} \arcsin \rho \tag{12}$$

$$\rho_S(\rho) = \frac{6}{\pi} \arcsin \frac{\rho}{2} \tag{13}$$

The t-Copula has a form

$$\boldsymbol{C}(u,v) = \boldsymbol{T}_{2,\rho,\nu}\{T_\nu^{-1}(u), T_\nu^{-1}(v)\}$$
$$= \int_{-\infty}^{T_\nu^{-1}(u)} \int_{-\infty}^{T_\nu^{-1}(v)} \frac{\Gamma\left(\frac{\nu+2}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\pi\nu\sqrt{1-\rho^2}} \tag{14}$$
$$\left(1 + \frac{s^2 - 2st\rho + t^2}{\nu}\right)^{-\frac{\nu+2}{2}} dsdt, \tag{15}$$

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

The copula density is

$$\boldsymbol{c}(u,v) = \frac{\boldsymbol{t}_{2,\rho,\nu}\{T_\nu^{-1}(u), T_\nu^{-1}(v)\}}{t_\nu\{T_\nu^{-1}(u)\} \cdot t_\nu\{T_\nu^{-1}(v)\}}, \tag{16}$$

where $\boldsymbol{t}_{2,\rho,\nu}$ is the pdf of $\boldsymbol{T}_{2,\rho,\nu}(\cdot,\cdot)$, and t_ν the density of standard t distribution.

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

2.2.2 Archimedean Copulae

The Archimedean copulae forms a large class of copulae with many convenient features. In general, they take a form

$$\boldsymbol{C}(u,v) = \psi^{-1}\{\psi(u), \psi(v)\}, \tag{17}$$

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.

The Frank copula (B3 in Joe (1997)) is a radial symmetric copula and cannot produce any tail dependence. It takes the form

$$\boldsymbol{C}_\theta(u,v) = \frac{1}{\theta}\log\left\{1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right\} \tag{18}$$

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

The Copula density is

$$\boldsymbol{c}_\theta(u,v) = \frac{\theta e^{\theta(u+v)(e^\theta - 1)}}{\left\{e^\theta - e^{\theta u} - e^{\theta v} + e^{\theta(u+v)}\right\}^2} \tag{19}$$

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

$$\tau_K(\theta) = 1 - 4\frac{D_1\{-\log(\theta)\}}{\log(\theta)}, \tag{20}$$

and

$$\rho_S(\theta) = 1 - 12\frac{D_2\{-\log(\theta)\} - D_1\{\log(\theta)\}}{\log(\theta)}, \tag{21}$$

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 (1997)) has upper tail dependence with the dependence parameter $\lambda^U = 2 - 2^{\frac{1}{\theta}}$ and displays no lower tail dependence.

$$\boldsymbol{C}_\theta(u, v) = \exp -\{(-\log(u))^\theta + (-\log(v))^\theta\}^{\frac{1}{\theta}}, \tag{22}$$

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.

$$\tau_K(\theta) = \frac{\theta - 1}{\theta} \tag{23}$$

The Clayton copula, by contrast to Gumbel copula, generates lower tail dependence in a form $\lambda^L = 2^{-\frac{1}{\theta}}$, but cannot generate upper tail dependence.

The Clayton copula takes the form

$$\boldsymbol{C}_\theta(u, v) = \left\{\max(u^{-\theta} + v^{-\theta} - 1, 0)\right\}^{-\frac{1}{\theta}}, \tag{24}$$

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

Kendall's τ to this copula dependency is

$$\tau_K(\theta) = \frac{\theta}{\theta + 2}. \tag{25}$$

2.2.3 Mixture Copula

Mixture copula is a linear combination of copulae. For a 2-dimensional random variable $\boldsymbol{X} = (X_1, X_2)^\top$, its distribution can be written as linear combination K copulae

$$\mathbf{P}(X_1 \le x_1, X_2 \le x_2) = \sum_{k=1}^{K} p^k \cdot \boldsymbol{C}^{(k)}\{F_{X_1}^{(k)}(x_1; \boldsymbol{\gamma}_1^{(k)}), F_{X_2}^{(k)}(x_2; \boldsymbol{\gamma}_2^{(k)}); \boldsymbol{\theta^{(k)}}\} \tag{26}$$

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 with the copula of the k^{th} component. The weights add up to one $\sum_{k=1}^{K} p^{(k)} = 1$.

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

$$\boldsymbol{C}(u, v) = \sum_{k=1}^{K} p^{(k)} \cdot \boldsymbol{C}^{(k)}\{F_{X_1}^{-1}(u), F_{X_2}^{-1}(v); \boldsymbol{\theta^{(k)}}\}. \tag{27}$$

The copula density is again a linear combination of copula density

$$\boldsymbol{c}(u,v) = \sum_{k=1}^{K} p^{(k)} \cdot \boldsymbol{c}^{(k)}\{F_{X_1}^{-1}(u), F_{X_2}^{-1}(v); \boldsymbol{\theta^{(k)}}\}. \tag{28}$$

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*

$$\rho_S = \sum_{k=1}^{K} p^{(k)} \cdot \rho_S^{(k)} \tag{29}$$

Proof. Spearman's ρ is defined as (Nelsen, 1999)

$$\rho_S = 12 \int_{\mathbb{I}^2} \boldsymbol{C}(s,t) ds dt - 3. \tag{30}$$

Rewrite the mixture copula into sumation of components

$$\rho_S = 12 \int_{\mathbb{I}^2} \sum_{k=1}^{K} p^{(k)} \cdot \boldsymbol{C}^{(k)}(s,t) ds dt - 3. \tag{31}$$

■

Example 4 *The Fréchet class can be seen as an example of mixture copula. It is a convex combinations of $\boldsymbol{W}$, $\boldsymbol{\Pi}$, and $\boldsymbol{M}$ (Nelsen, 1999)*

$$\boldsymbol{C}_{\alpha,\beta}(u,v) = \alpha \boldsymbol{M}(u,v) + (1-\alpha-\beta)\boldsymbol{\Pi}(u,v) + \beta \boldsymbol{W}(u,v), \tag{32}$$

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

$$\tau_K(\alpha,\beta) = \frac{(\alpha-\beta)(\alpha+\beta+2)}{3} \tag{33}$$

, and

$$\rho_S(\alpha,\beta) = \alpha - \beta \tag{34}$$

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

$$\boldsymbol{C}(u,v) = p \cdot \boldsymbol{C}^{\text{Gaussian}}(u,v) + (1-p)(uv). \tag{35}$$

The corresponding copula density is

$$\boldsymbol{c}(u,v) = p \cdot \boldsymbol{c}^{\text{Gaussian}}(u,v) + (1-p). \tag{36}$$

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, so as the optimal hedging ratio h^*. 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.

2.3 Other Copula

The Plackett copula has an expression

$$\boldsymbol{C}_\theta(u,v) = \frac{1+(\theta-1)(u+v)}{2(\theta-1)} - \frac{\sqrt{\{1+(\theta-1)(u+v)\}^2 - 4uv\theta(\theta-1)}}{2(\theta-1)} \tag{37}$$

$$\rho_S(\theta) = \frac{\theta+1}{\theta-1} - \frac{2\theta\log\theta}{(\theta-1)^2} \tag{38}$$

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

$$\begin{aligned} &\frac{\mathbf{P}(U\le u, V\le v)\cdot \mathbf{P}(U>u, V>v)}{\mathbf{P}(U\le u, V>v)\cdot \mathbf{P}(U>u, V\le v)} \\ &= \frac{\boldsymbol{C}_\theta(u,v)\{1-u-v+\boldsymbol{C}_\theta(u,v)\}}{\{u-\boldsymbol{C}_\theta(u,v)\}\{v-\boldsymbol{C}_\theta(u,v)} \\ &= \theta. \end{aligned} \tag{39}$$

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 Estimation

3.1 Simulated Method of Moments

This method is suggested by Oh and Patton (2013). In our 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

$$\rho_S = 12\int\int_{I^2} C_{\boldsymbol{\theta}}(u,v)\,\mathrm{d}u\,\mathrm{d}v - 3 \tag{40}$$

$$\tau_K = 4\,\mathsf{E}[C_{\boldsymbol{\theta}}\{F_X(x),F_Y(y)\}] - 1, \tag{41}$$

$$\lambda_q = \begin{cases} \mathbf{P}(F_X(X)\le q|F_Y(Y)\le q) = \dfrac{C_{\boldsymbol{\theta}}(q,q)}{q}, & \text{if } q\in(0,0.5], \\ \mathbf{P}(F_X(X)>q|F_Y(Y)>q) = \dfrac{1-2q+C_{\boldsymbol{\theta}}(q,q)}{1-q}, & \text{if } q\in(0.5,1). \end{cases} \tag{42}$$

The empirical counterparts are

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

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

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

$$
\hat{\boldsymbol{\theta}} = \operatorname*{argmin}_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \boldsymbol{g}(\boldsymbol{\theta})^\top \hat{\boldsymbol{W}} \boldsymbol{g}(\boldsymbol{\theta}),
$$

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

In this work, we use $\tilde{\boldsymbol{m}}(\boldsymbol{\theta}) = (\rho_S, \lambda_{0.05}, \lambda_{0.1}, \lambda_{0.9}, \lambda_{0.95})^\top$ 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

$$
\boldsymbol{f}_{\boldsymbol{X}}(x_1, ..., x_d) = \boldsymbol{c}\{F_{X_1}(x_1), ..., F_{X_d}(x_d)\} \prod_{j=1}^{d} f_{X_i}(x_i). \tag{43}
$$

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 $\sum_{i=1}^{n} \boldsymbol{f}_{\boldsymbol{X}}(X_{i,1}, ..., X_{i,d})$ can be decomposed into dependence part and marginal part,

$$
\begin{aligned}
L(\boldsymbol{\theta}) &= \sum_{i=1}^{n} \boldsymbol{c}\{F_{X_1}(x_{i,1}; \boldsymbol{\delta}_1), ..., F_{X_d}(x_{i,d}; \boldsymbol{\delta}_d); \boldsymbol{\gamma}\} + \sum_{i=1}^{n}\sum_{j=1}^{d} f_{X_j}(x_{i,j}; \boldsymbol{\delta}_j) && (44) \\
&= L_C(\boldsymbol{\delta}_1, ..., \boldsymbol{\delta}_d, \boldsymbol{\gamma}) + \sum_{j=1}^{d} L_j(\boldsymbol{\delta}_j) && (45)
\end{aligned}
$$

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

$$\left(\frac{\partial L_1}{\partial \boldsymbol{\delta}_1}, ..., \frac{\partial L_d}{\partial \boldsymbol{\delta}_d}, \frac{\partial L_C}{\partial \boldsymbol{\gamma}} \right) = \mathbf{0}. \tag{46}$$

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

$$\hat{\boldsymbol{\gamma}} = \underset{\boldsymbol{\gamma}}{\operatorname{argmax}} \sum_{i=1}^{n} \boldsymbol{c}\{\hat{F}_{X_1}(x_{i,1}), ..., \hat{F}_{X_d}(x_{i,d}); \boldsymbol{\gamma}\}. \tag{47}$$

3.3 Comparison

Both the simulated method of moments and the maximum likelihood estimation are unbiased. The problem remain is which procedure is more suitable for hedging.

Figure 1 shows the empirical quantile dependence of Bitcoin and CME future and the copula implied quantile dependence from MLE and MM calibration procedures. Although the MLE is a better fit to a range of quantile dependence in the middle, it fails to address the situation in the tails. Our data empirically has weaker quantile dependence in the ends, and those points generate PnL to the hedged portfolio. MM is preferred visually as it produces a better fit to the dependence structure in the two extremes.

4 Results

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

[The issue with the Frank copula is that is has no tails. A scatterplot looks like a strip, there is no concentration in the tails. For CDO pricing (and this is what I remember from my PhD studies) this poses problems as you move from senior to junior tranches. Here, I suppose it just does not capture the empirical behaviour of the data.]

4.1 Hedging Effectiveness

The hedging effectiveness (HE) is defined as

$$1 - \frac{\rho(r^h)}{\rho(r^S)}. \tag{48}$$

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

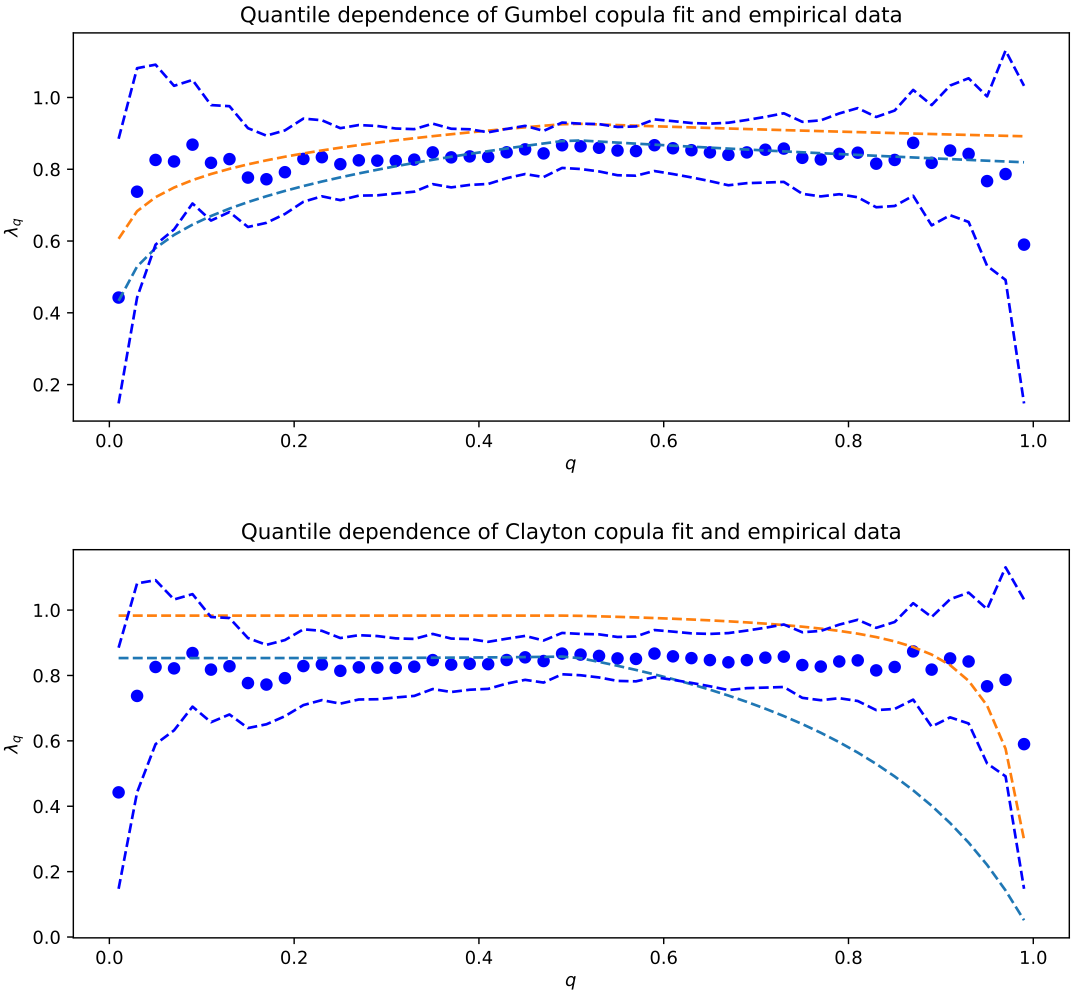


Figure 1: Quantile dependences of Gumbel, and Clayton Copula. The blue circle dots are the quantile dependence estimate of Bitcoin and CME future, blue dotted lines are the estimates' 90% confidnece interval. Orange dotted line is the copula implied quantile dependence by MM estimation. Light blue dotted line is the copula implied quantile dependence by MM estimation.

in section 2.1, e.g. via Expected Shortfall (ES)

$$1 - \frac{\mathrm{ES}_\alpha(r^h)}{\mathrm{ES}_\alpha(r^S)}. \tag{49}$$

The box-plots in figure 6 show the out-of-sample hedging effectiveness of different copulas under various risk reduction objectives across testing datasets. Observe that most of the copulae perform well. 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), Giacomini et al. (2009), and also Manner and Reznikova (2012).

Frank-copula, in general, is not a good choice to model financial data. We can see from figure 2 that the Frank copula is not fitting the Bitcoin and its future visually, no matter which optimization procedure is being deployed. The samples of Frank diffuse like a strip with parallel edge when the parameter θ decrease (samples are being less dependent). This makes Frank copula not a a good fit to the data.

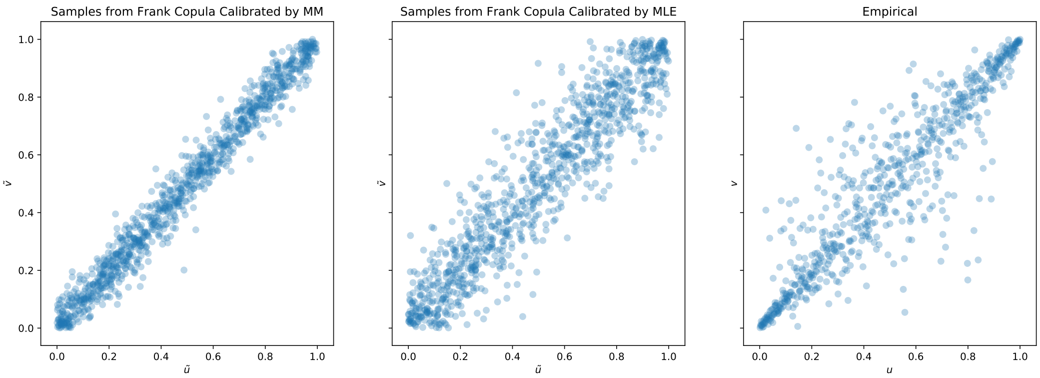


Figure 2: Comparison of Frank Copula Samples and Pseudo Observations of Bitcoin and CME Future Returns.

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 4 tabulates the time series of out-of-sample returns of hedged portfolio under various copulas and risk reduction objectives.

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 3: First Panel: Out of Sample Log-return of Bitcoin; Second Panel: Out of Sample Log-return of Future; Third Panel: Out of Sample Log-return of Hedged Portfolio by Gumbel copula with the aim of variance reduction. The red dots indicate the lowest 10% return of Bitcoin, i.e. large negative moments of price. Forth Panel: Out of Sample Log-return of Hedged Portfolio by $h = 1$ (naive hedge).

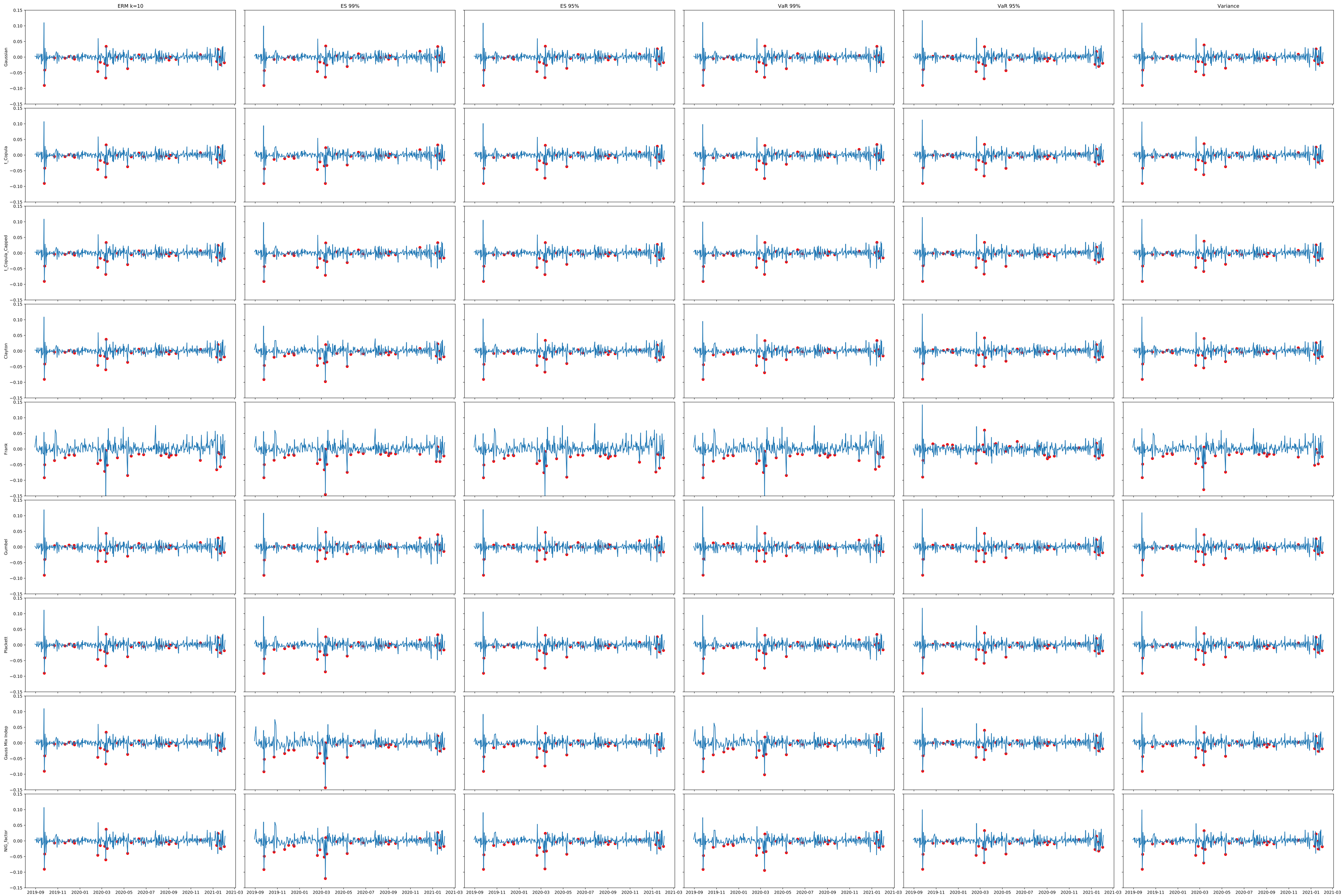


Figure 4: Out-of-Sample Returns of Hedged Portfolio of Copulas and Risk Reduction Objectives.

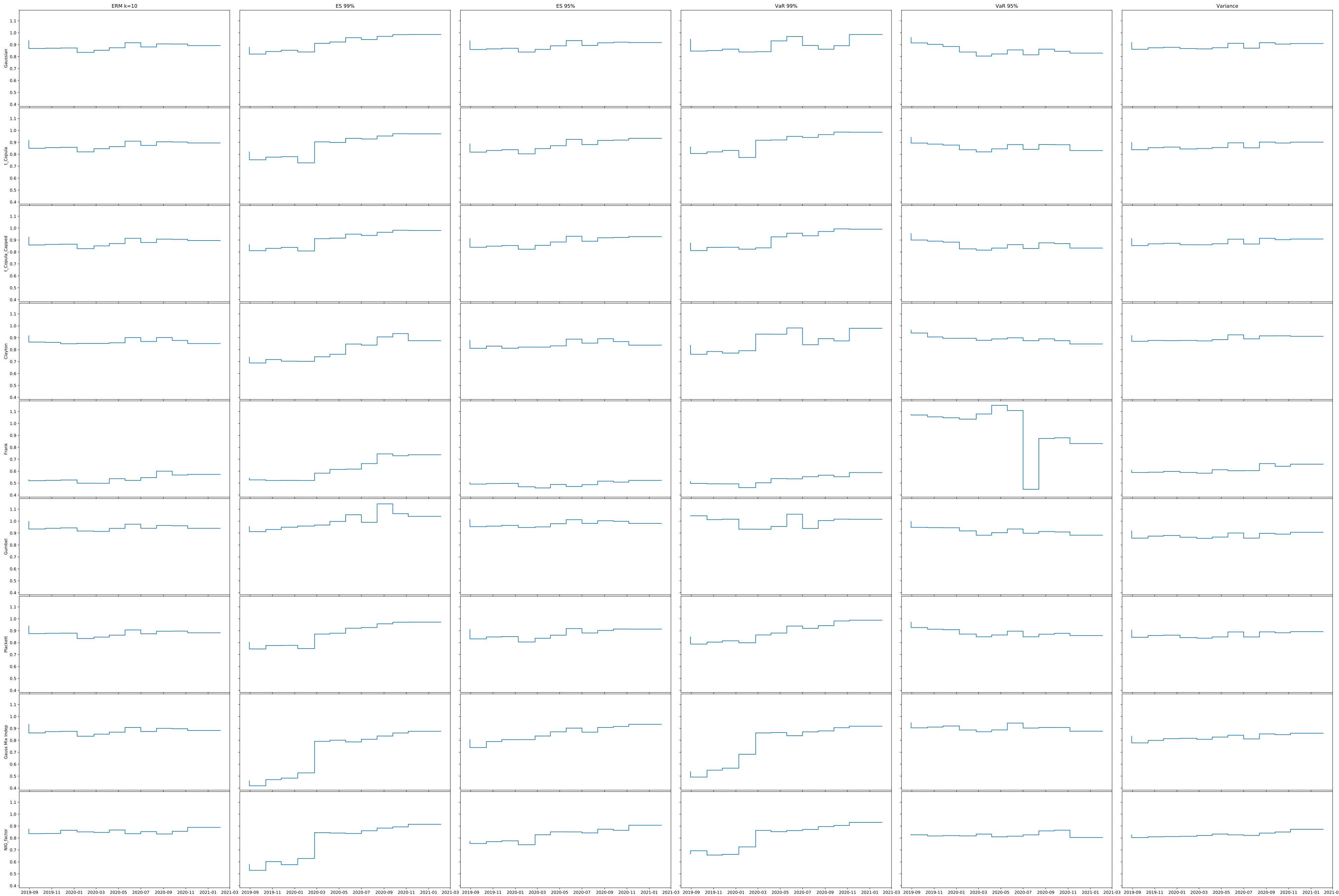


Figure 5: Optimal Hedge Ratio Obtained from Combinations of Copula and Risk Reduction Objective.

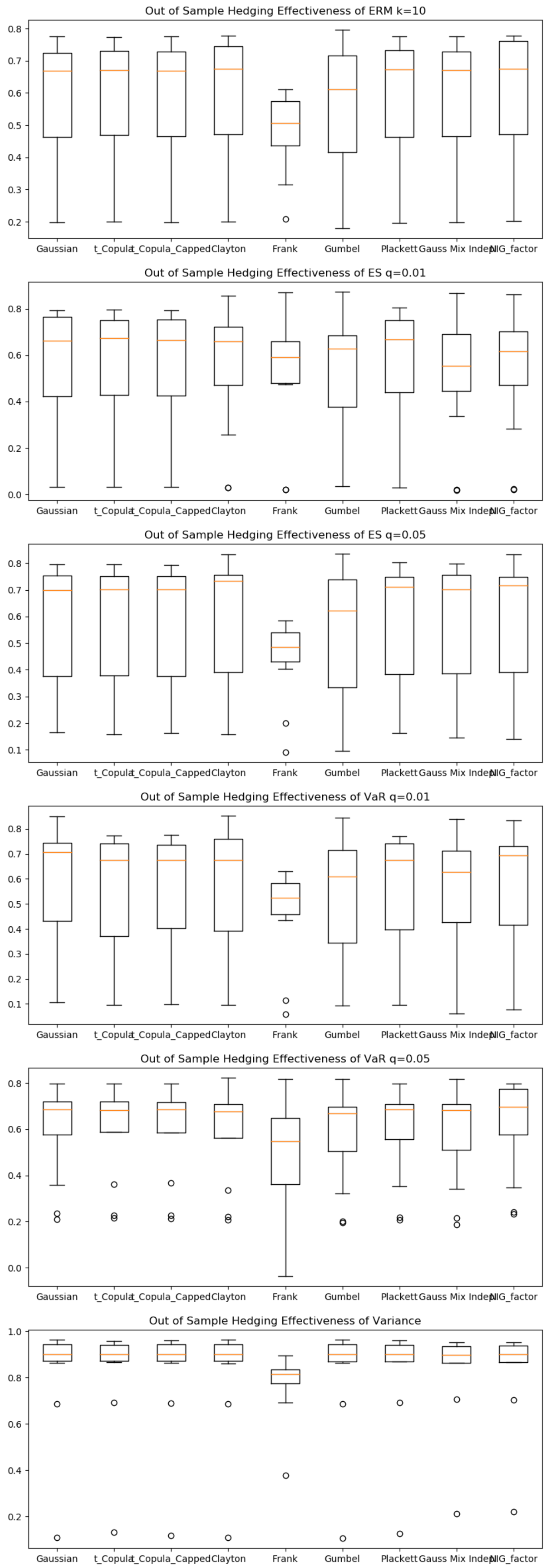


Figure 6: Out of Sample Hedging Effectiveness Box-plot. The HEs are obtained from a set of out-of-sample data, each set consists 30 days log returns of Bitcoin and CME future.

Figure 3 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.

4.2 Robustness

The study of robustness concerns the stability of statistical estimation with respect to violation in assumptions. In our context, the robustness is with respect to outliers (or jumps). It is natural to do we want the optimal hedge ratio react to extreme market changes? In practice, outliers of returns can come from anywhere, for example, a tweet from Elon Musk, a sudden large order from institutional investor, or an incident of system failure in cryptocurrency exchanges. Rapid and drastic changes in portfolio weight causes problem of slippage and transaction cost. Investors should be aware of the cost brought by the sensitivity of the optimal hedge ratio procedure.

The discussion of sensitivity or robustness dates back to Huber and Ronchetti (1981)'s work on robust statistics. Hampel et al. (2011) suggest an infinitesimal approach to investigate sensitivity of statistical procedures. There are three central concepts in this approach, qualitative robustness, influence function, and break-down point. They are loosely related to the concept of continuity, first derivative of functional, and the distance of a functional to its nearest pole (singularity). While the first concept is a qualitative feature of a functional, the second the third concepts are practical tools to measure sensitivity quantitatively. We deploy a finite sample version of the second and third concepts. Details of robustness of risk measures can be found in Cont et al. (2010).

The influence function of $\hat{h}_\rho$ with finite sample size n is

$$\mathrm{IF}(\boldsymbol{z};\hat{h}_\rho) = \frac{\hat{h}_\rho(\boldsymbol{X}_1,...,\boldsymbol{X}_n,\boldsymbol{z}) - \hat{h}_\rho(\boldsymbol{X}_1,...,\boldsymbol{X}_n)}{\frac{1}{n+1}}. \tag{50}$$

[The inclusion of $\boldsymbol{z}$ has nothing to do with the probability in a probability space, i.e. it is possible to include points with density zero.]

The equation describes the effect of a single contamination at point $\boldsymbol{z}$ on the estimate of OHR, standardised by the mass of the contamination.

Figure 7 shows the influence function of $\hat{h}_\rho$ of using t copula estimated by MLE with 300 data points of Bitcoin and CME future returns from 14/12/2018 to 25/02/2020. Contamination are in a set $\{-0.3,-0.27,...,0.3\}\times\{-0.3,-0.27,...,0.3\}$, in total 900 pairs of contamination. The product is Cartesian product of two sets.

We can see from the plots that Expected Shortfall with $\alpha = 99\%$ is very sensitive the negative return in spot (lower right plot). The h^* obtained this way increases with a single contamination of negative jump in spot price. VaR at 99% is also sensitive to negative jump in spot price but with a lower level (lower left plot). This is a natural result that reflects investor's strong preference on risk avoidance: investor increases her future's short position to compensate a large drop in spot price she saw in her data. The result of ES being more sensitive to VaR as risk measure agrees with the conclusion of Cont et al. (2010).

Other risk measures are relatively less sensitive. Interestingly, although ERM places heavy weights to negative returns, its IF is similar to that of variance, where variance does not exhibit risk preference. [FL:This might due to the smooth $\phi(p)$ over the spectrum $[0,1]$ of ERM. The $\phi(p)$ of VaR is a Dirac function at a single point α, that of ES has a sharp cut off at $1-\alpha$, a tiny change in rank of r^h (caused by a contamination $\boldsymbol{z}$) causes VaR and ES to shift their weights.]

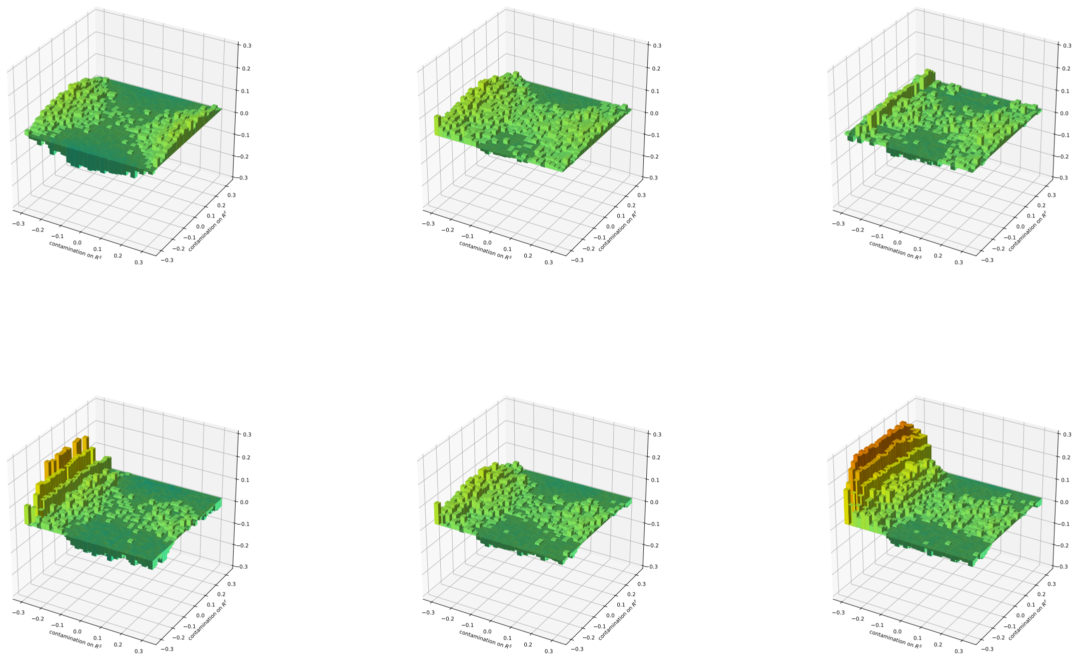


Figure 7: Influence functions (IF) of h^* using t copula copula estimated by MLE. From left to right, top to bottem, the plots are IF of using Var, ERM_{10}, $\mathrm{VaR}_{0.95}$, $\mathrm{VaR}_{0.99}$, $\mathrm{ES}_{0.95}$, and $\mathrm{ES}_{0.99}$ respectively.

References

Acerbi, C. (2002): "Spectral measures of risk: A coherent representation of subjective risk aversion," *Journal of Banking & Finance*, 26, 1505–1518.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999): "Coherent measures of risk," *Mathematical finance*, 9, 203–228.

Barbi, M. and S. Romagnoli (2014): "A Copula-Based Quantile Risk Measure Approach to Estimate the Optimal Hedge Ratio," *Journal of Futures Markets*, 34, 658–675.

Bollerslev, T., V. Todorov, and L. Xu (2015): "Tail risk premia and return predictability," *Journal of Financial Economics*, 118, 113–134.

Cherubini, U., S. Mulinacci, and S. Romagnoli (2011): "A copula-based model of speculative price dynamics in discrete time," *Journal of Multivariate Analysis*, 102, 1047–1063.

Cont, R., R. Deguest, and G. Scandolo (2010): "Robustness and sensitivity analysis of risk measurement procedures," *Quantitative finance*, 10, 593–606.

Creal, D., S. J. Koopman, and A. Lucas (2008): "A general framework for observation driven time-varying parameter models," Tech. rep., Tinbergen Institute Discussion paper.

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

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

Ederington, L. H. and J. M. Salas (2008): "Minimum variance hedging when spot price changes are partially predictable," *Journal of Banking & Finance*, 32, 654–663.

Embrechts, P., A. McNeil, and D. Straumann (2002): "Correlation and dependence in risk management: properties and pitfalls," *Risk management: value at risk and beyond*, 1, 176–223.

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

Fama, E. F. (1963): "Mandelbrot and the stable Paretian hypothesis," *The journal of business*, 36, 420–429.

Fisher, N. I. and P. K. Sen (2012): *The collected works of Wassily Hoeffding*, Springer Science & Business Media.

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

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

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

Hampel, F. R., E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel (2011): *Robust statistics: the approach based on influence functions*, vol. 196, John Wiley & Sons.

Härdle, W. and S. Léopold (2011): *Applied Multivariate Statistical Analysis*, Springer, 4th ed.

Härdle, W. K., N. Hautsch, and L. Overbeck (2008): *Applied quantitative finance*, Springer Science & Business Media.

Härdle, W. K. and O. Okhrin (2010): "De copulis non est disputandum," *AStA Advances in Statistical Analysis*, 94, 1–31.

Hoeffding, W. (1940): "Scale-invariant correlation theory (English translation)," 5, 181–233.

——— (1941): "Scale-invariant correlations for discontinuous distributions (English translation)," 7, 49–70.

Huber, P. and E. Ronchetti (1981): "Robust statistics, ser," *Wiley Series in Probability and Mathematical Statistics. New York, NY, USA, Wiley-IEEE*, 52, 54.

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

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

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

Nelsen, R. B. (1999): *An Introduction to Copulas*, Springer.

Oet, M. V., J. Dooley, D. Gramlich, P. Sarlin, and S. J. Ong (2015): "Evaluating the Information Value for Measures of Systemic Conditions," .

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

Schweizer, B., E. F. Wolff, et al. (1981): "On nonparametric measures of dependence for random variables," *Annals of Statistics*, 9, 879–885.

Sklar, A. (1959): "Fonctions de répartition a n dimensions et leurs marges," *Publications de l'Institut de Statistique de l'Universite de Paris*, 8, 229–231.

Trimborn, S. and W. K. Härdle (2018): "CRIX an Index for cryptocurrencies," *Journal of Empirical Finance*, 49, 107–122.

5 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, pdf of the linear combination of marginals $Z = n_1 \cdot X_1 + ... + n_d \cdot X_d$ is*

$$f_Z(z) = \left|n_1^{-1}\right| \int_{[0,1]^{d-1}} \boldsymbol{c}_{X_1,...,X_d}\{F_{X_1} \circ S(z), u_2, ..., u_d\} \cdot f_{X_1} \circ S(z) du_2...du_d \tag{51}$$

$$S(z) = \frac{1}{n_1} \cdot z - \frac{n_2}{n_1} \cdot F^{(-1)}_{X_2}(u_2) - ... - \frac{n_d}{n_1} \cdot F^{(-1)}_{X_d}(u_d) \tag{52}$$

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

$$\begin{bmatrix} Z \\ X_2 \\ \vdots \\ X_d \end{bmatrix} = \begin{bmatrix} n_1 & n_2 & \cdots & n_d \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & & 1 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_d \end{bmatrix}. \tag{53}$$

By transformation variables

$$\boldsymbol{f}_{Z,X_2,...,X_d}(z, x_2, ..., x_d) = \boldsymbol{f}_{X_1,...,X_d}\left(\boldsymbol{A}^{-1} \begin{bmatrix} z \\ x_2 \\ \vdots \\ x_d \end{bmatrix}\right) \cdot |\det \boldsymbol{A}^{-1}| \tag{54}$$

$$= \left|n_1^{-1}\right| \boldsymbol{f}_{X_1,...,X_d}\{S(z), x_2, ..., x_d\} \tag{55}$$

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

$$\boldsymbol{c}_{X_1,...,X_d}(u_1, ..., u_d) = \frac{\boldsymbol{f}_{X_1,...,X_d}(x_1, ..., x_d)}{\prod_{i=1}^d f_{X_i}(x_i)}, \tag{56}$$

we have

$$\boldsymbol{f}_{Z,X_2,...,X_d}(z, x_2, ..., x_d) = \tag{57}$$

$$\left|n_1^{-1}\right| \cdot \boldsymbol{c}_{X_1,...,X_d}\{F_{X_1} \circ S(z), u_2, ..., u_d\} \cdot f_{X_1}\{S(z)\} \cdot \prod_{i=2}^d f_{X_i}(x_i) \tag{58}$$

The claim 51 is obtained by integrating out $x_2, ...x_d$ by substituting $dx_i = \frac{1}{f_{X_i}(x_i)} du_i$. ■

6 Data

This section is under construction Cryptocurrenices are traded around the clock, but CME future are traded from Sunday to Friday from 05:00 p.m. to 04:00 p.m. U.S. central time. We match the timestamps and timezones of different data sources.

#	Asset	Data Source	Type	Tradable at CT[1]	Tradable at CET[2] during CST[3]	Tradable at CET during CDT[4]	Tradable at UTC during CST	Tradable at UTC during CDT
1	Bitcoin	Coingecko API	Hourly Close		11:00pm D+0	11:00pm D+0	10:00pm D+0*	10:00pm D+0*
2	CME Future	Bloomberg	Daily Open	05:00pm D-1	00:00am D+0*	00:00am D+0*	11:00pm D-1	10:00pm D-1
3	CME Future	Bloomberg	Daily Close	04:00pm D+0	11:00pm D+0*	11:00pm D+0*	10:00pm D+0	09:00pm D+0
4	CRIX	IRTG (from Coingecko)	Index					00:00am D+0*

Table 1: * indicates the timestamp of raw data from data source.

Hedging Pair 1 is hedging #1 (Bitcoin Spot) with #3 (CME future). The time difference between the two prices is zero. They are both adjusted to CET time: #1 by pandas.Series.dt.tz_convert; #3 by retrieving data from Bloomberg Terminal located in Berlin.

Hedging Pair 2 is hedging #4 (CRIX) with #2 (CME future). We observe #2 two hours and one hour before #4 during CST and CDT respectively.

6.1 Time Difference

[1] CT stands for U.S. Central Time. It represents two observances of time, the Central Standard Time (CST) and the Central Daylight Time (CDT)

[2] CET stands for Central European Time. It is one hour ahead UTC.

[3] CST is six hours behind UTC.

[4] CDT is five hours behind UTC.

	Open	High	Low	Close
2021-02-02 23:00	36360.0	38155.0	36240.0	37790.0
2021-02-01 23:00	34205.0	36665.0	34070.0	36535.0
2021-01-31 23:00	33715.0	35280.0	32800.0	34265.0
2021-01-28 23:00	33995.0	39530.0	32590.0	35180.0
2021-01-27 23:00	31005.0	33710.0	30350.0	33085.0

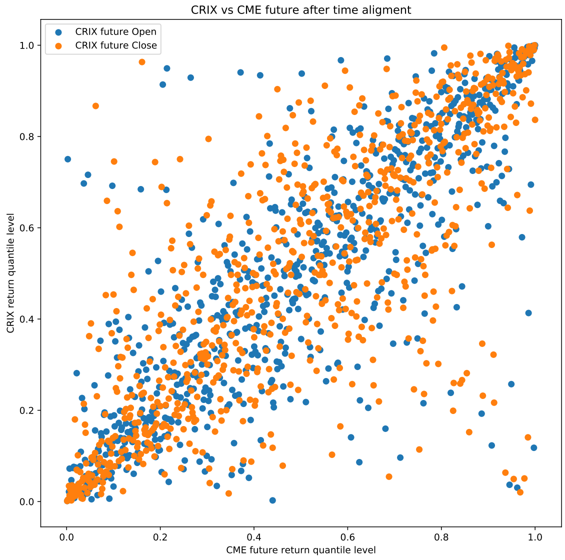
Table 2: CME Bitcoin Future Raw Data

	date	CRIX	future	log return CRIX	log return future
0	2021-02-04	104518.468839	38080.0	0.054757	0.046220
1	2021-02-03	98949.179255	36360.0	0.059741	0.061097
2	2021-02-02	93210.948461	34205.0	0.002204	0.014429
3	2021-02-01	93005.711051	33715.0	0.013628	-0.008271
4	2021-01-29	91746.863103	33995.0	0.081917	0.092065

Table 3: CRIX #4 with Opening price of CME Bitcoin future #2 and their log returns

	date	CRIX	future	log return CRIX	log return future
0	2021-02-05	103348.488555	38220.0	-0.011257	0.011314
1	2021-02-04	104518.468839	37790.0	0.054757	0.033774
2	2021-02-03	98949.179255	36535.0	0.059741	0.064146
3	2021-02-02	93210.948461	34265.0	-0.016175	-0.026353
4	2021-01-30	94730.919657	35180.0	0.032007	0.061398

Table 4: CRIX #4 with Closing price of CME Bitcoin future #3 shifted for one day (D-1) and their log returns



Kendall's tau between CRIX and future Close is 0.608429;
Kendall's tau between CRIX and future Open is 0.673266; we pick this unless we have hourly CRIX.

6.2 Statistics of Percentage Difference Between CME Bitcoin future Open Price and Last Close Price

$$\text{diff} = \frac{\text{Open}_t - \text{Close}_{t-1}}{\text{Close}_{t-1}}$$

Mean of diff = 0.00236
Std of diff = 0.02206
Max of diff = 0.16394
UQ of diff = 0.00814
Median of diff = 0.00132
LQ of diff = -0.00412
Min of diff = -0.12190