



Discovering Intraday Tail Dependence Patterns via a Full-Range Tail Dependence Copula

Lei Hua 匝

Article

Department of Statistics and Actuarial Science, Northern Illinois University, DeKalb, IL 60115, USA; lhua@niu.edu

Abstract: In this research, we employ a full-range tail dependence copula to capture the intraday dynamic tail dependence patterns of 30 s log returns among stocks in the US market in the year of 2020, when the market experienced a significant sell-off and a rally thereafter. We also introduce a model-based unified tail dependence measure to directly model and compare various tail dependence patterns. Using regression analysis of the upper and lower tail dependence simultaneously, we have identified some interesting intraday tail dependence patterns, such as interactions between the upper and lower tail dependence over time among growth and value stocks and in different market regimes. Our results indicate that tail dependence tends to peak towards the end of the regular trading hours, and, counter-intuitively, upper tail dependence tends to be stronger than lower tail dependence for short-term returns during a market sell-off. Furthermore, we investigate how the Fama–French five factors affect the intraday tail dependence patterns and provide plausible explanations for the occurrence of these phenomena. Among these five factors, the market excess return plays the most important role, and our study suggests that when there is a moderate positive excess return, both the upper and lower tails tend to reach their lowest dependence levels.

Keywords: unified tail dependence measure; PPPP copula; Fama–French five factors; regression on tail dependence; multiple components GARCH

1. Introduction

Dependence modeling among financial assets has been an important topic of research in quantitative risk management. For daily returns, there are many papers on the use of copulas to model dependence structures across different financial assets, such as Dißmann et al. (2013) and Krupskii and Joe (2020), among many others. However, for intraday financial risks, there has not yet been much work conducted. The recent market crash on 16 March 2020 tells us that intraday risks of the financial market can be as large as a loss of 12% in the major stock market index in a single day. With the growing availability of high-frequency financial data and the massive popularity of algorithm trading, it becomes more and more relevant to study intraday risks of the financial market. In particular, little is known about the intraday tail dependence patterns of financial assets and these are more relevant to better understand the intraday market operation and tail risks.

There has been a research trend on dependence modeling for high-frequency financial data. For example, Koopman et al. (2018) used copulas with time-varying parameters to show that in the US stock markets, the dependence starts low but gradually increases throughout the day. The copula was used by Carvalho et al. (2021) to investigate the dependence of the 15 min returns of the Brazilian stock market. Buccheri et al. (2021) investigated intraday dependence patterns in asynchronous high-frequency financial data. Kim and Hwang (2021) studied directional dependence of intraday volatility via the copulas. Weiß and Supper (2013) utilized vine copulas to find tail dependence between intraday returns and bid–ask spreads of Nasdaq stocks.

To our knowledge, there is lack of research work in the literature about tail dependence patterns of high-frequency financial returns, how they change over time between different



Citation: Hua, Lei. 2023. Discovering Intraday Tail Dependence Patterns via a Full-Range Tail Dependence Copula. *Risks* 11: 195. https:// doi.org/10.3390/risks11110195

Academic Editor: Mogens Steffensen

Received: 21 October 2023 Revised: 4 November 2023 Accepted: 7 November 2023 Published: 11 November 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). stock categories and in different market regimes, and what factors contribute to such changes. From what we have learned in this study, intraday dependence patterns can be quite different than those in daily returns and they are worthy of further study.

Copula models such as Student-*t*, Gumbel, Clayton, and BB series copulas such as BB1, BB4, and BB7 copulas studied in Joe (1997) are commonly used to model dependence among returns of financial assets, as those copulas can capture tail dependence in financial markets. For example, Carvalho et al. (2021) used the BB7 copula to model asymmetric tail dependence between financial returns. However, the commonly used copula models are not flexible enough in the tails to serve our purpose of modeling tail dependence, especially for conducting tail comparisons. The Student-*t* copula, for example, is only symmetric between the upper and lower tails and can only handle relatively stronger tail dependence; the other commonly used copulas, such as the BB series copulas, cannot be symmetric between the upper and lower tails and a direct comparison between the upper and lower tails using such copulas is unfeasible.

To capture various intraday tail dependence patterns, one needs not only a model to capture the overall dependence but also a model that is tailored to be flexible enough to model the dynamic lower and upper tails and make comparisons between the tails without too much distraction. The full-range tail dependence copulas studied in Hua (2017) and Su and Hua (2017) are ideal candidate models for this study, due to their flexible upper and lower tails to capture various tail dependence patterns. In particular, we use the PPPP copula studied in Su and Hua (2017) to model the dependence between standardized residuals of high-frequency log returns of two financial assets, with each univariate margin being modeled by a multiple components GARCH model to account for intraday seasonality, volatility, etc. The two dependence parameters that are used to respectively control the lower and upper tails of the PPPP copula are linked to a generalized additive model with natural cubic splines being utilized to explain the nonlinear effects arising from the time and the Fama–French five factors. To quantify the strength of tail dependence for comparisons, we propose a model-based unified tail dependence measure based on the PPPP copula. The unified measure ranges from 0 to 1 and can quantify the full range of positive tail dependence, covering both asymptotic dependence and asymptotic independence cases.

Through the regression analyses mentioned above, we have observed that, for QQQ and SPY, which are ETFs that track the Nasdaq-100 and S&P 500 indexes, respectively, the unified upper and lower tail dependence increases over time and tends to peak toward 4 p.m. in New York time, which is the end of the regular trading hours of the US stock market, and the tail dependence often plateaus or drops slightly during noon; this phenomenon also exists between AAPL and MSFT, the two main growth stocks. For the value stocks, we studied the pair of two major banks, JPM and BAC, for which the tail dependence can be higher after the market opens and before the market closes, with the noon trading session having the lowest unified tail dependence measures.

Market regimes also affect the overall dependence of the tail and the degree of asymmetry between the dependence of the lower and upper tails. It is commonly believed that for the returns of financial assets, the lower tail dependence is stronger than the upper tail dependence, and the lower tail dependence is more relevant when the stock market crashes due to some spillover or contagious effects (Kato et al. 2022; Rodriguez 2007). On the contrary, we find that high-frequency returns tend to have higher upper tail dependence than lower tail dependence, especially during a market sell-off, when both the upper and lower tail dependence increases, and furthermore, the lower and upper tail dependence patterns become a lot more different, much more than those during a market rally.

Among the Fama–French five factors (Fama and French 2015), the factor of excess return on the market, which is the difference between value-weighted market returns and a risk-free return rate, affects the tail dependence patterns the most. Both the lower and upper tail dependence reaches the minimum when there is a moderate positive excess return. Some other factors such as "Small Minus Big" and "Conservative Minus Aggressive"

also affect tail dependence significantly, with the exception that the "Robust Minus Weak" factor was not significant after the other four factors were controlled in our analyses.

We believe that our work is the first one to use a full-range tail dependence copula to study the tail dependence patterns of short-term financial asset returns. The proposed method of using a full-range tail dependence copula with regression on both the upper and lower tail dependence parameters provides an effective way of modeling tail dependence and their explanatory variables. The model-based unified tail dependence measure overcomes the difficulty of comparing tail dependence with different quantities in different situations. With the help of the proposed model, we have obtained some new empirical findings on tail dependence patterns for high-frequency financial data.

In what follows, we discuss the details with the following structures: Section 2 explains the details of the regression models and covers the notion of full-range tail dependence copulas, the unified tail dependence measures, etc. Section 3 presents an empirical study for some US stocks and ETFs that track the major stock indexes, and the main findings on intraday tail dependence patterns will be discussed in detail. Finally, we will conclude the paper in Section 4.

2. Regression Models for Intraday Tail Dependence

In Section 2.1, we will briefly describe how to use the copula to model the dependence between several time series data. The concept of the *full-range tail dependence copula* and its properties will be introduced in Section 2.2. To directly compare the upper and lower tail dependence, we propose using a model-based tail dependence measure referred to as the *unified tail dependence measure*, which is discussed in Section 2.3. Section 2.4 introduces a GARCH model of multiple components and the model is used to account for intraday features of returns, such as seasonality and dynamic volatilities. Section 2.5 discusses the generalized additive models for the regression analysis of dependence parameters.

To facilitate the reading, we provide in Figure 1 a flow chart of the proposed models and their associated sections.



Figure 1. Model structures and their corresponding sections. "TS1" and "TS2" are "Time series of log returns of the first ticker" and "Time series of log returns of the second ticker", respectively, "Expl. Var." is "Explanatory Variables", "Std. Res." is "Standardized Residuals", and "Unif. Tail Dep." is "Unified Tail Dependence".

2.1. Copula for Dependence of Multiple Time Series

Sklar's theorem (Sklar 1959) tells us that for two continuous random variables *X* and *Y* with the joint CDF *F* and the marginal CDFs F_X and F_Y , respectively, the copula function $C(u_x, u_y) := F(F_X^{-1}(u_x), F_Y^{-1}(u_y))$ characterizes the dependence structure between *X* and *Y*. We refer to Joe (1997) for basic concepts about copulas and various parametric copula functions.

Copula functions, unlike many other dependence measures in quantitative finance, such as covariances and correlations, have the advantage of capturing non-linear dependence structures, regardless of how flexible the dependence patterns are; see, for example, Patton (2009). It is especially useful to capture the dependence between financial asset

returns, which is generally higher during a market sell-off or rally; see De Luca and Zuccolotto (2017). Copula functions can be utilized to capture dependence in the upper and lower tails for such stronger dependence structures that can be caused by greed and fear of the market participants.

Copula functions are frequently used to model residuals that arise from each marginal time series, which are usually modeled individually by their appropriate models. In Patton (2012), for example, log returns of each financial asset are represented using commonly known time series models to account for mean and variance processes, and copula functions are utilized to model residual dependence. The dependence structures that are modeled by copulas are actually the dependence structures of some "deviance" from their expected values after accounting for the past information on expectations and volatilities, and the dependence suggested by the copula fitted based on the standardized residuals reflects how new information affects both the marginal time series. We use the copula to model such dependence structures between time series, but it should not be simply understood as the dependence between the observed time series data, and it only explains how marginal time series depend on each other after new information arrives. For the intraday financial data, we also use the copula to model their standardized residuals. To account for dynamic intraday volatility patterns, we employ multiple component GARCH models, and ARMA models are applied for the mean processes of log returns. Then, standardized residuals are utilized to fit the copulas; see Section 2.4 for details.

2.2. Full-Range Tail Dependence Copula

For the copula *C*, if $\lim_{u\to 0^+} C(u,u)/u = \lambda_L \in (0,1]$, then λ_L is referred to as the lower tail dependence parameter of *C*. Let $\widehat{C}(u_1, u_2) := C(1 - u_1, 1 - u_2) + u_1 + u_2 - 1$ be the survival copula of *C*. If $\lim_{u\to 0^+} \widehat{C}(u, u)/u = \lambda_U \in (0,1]$, then λ_U is referred to as the upper tail dependence parameter for *C*. In the literature, the tail dependence parameters are also known as tail dependence coefficients and they can be used to describe the degree of dependence in either the upper or lower tails. If we are concerned about the lower tail, which corresponds to the dependence of high-quantile losses between financial assets, and $C(u, u) \approx \lambda_L u$ when u is a small positive number very close to 0, then we can approximate the probability $\mathbb{P}[X_1 \leq F_{X_1}^{-1}(u), X_2 \leq F_{X_2}^{-1}(u)]$ by $\lambda_L u$, where X_1 and X_2 are random variables and $F_{X_1}^{-1}$ and $F_{X_2}^{-1}$ are the corresponding quantile functions. In practice, parametric copulas are frequently used to fit real data, and the estimated values of λ_L and λ_U can be used to assess risk and profit dependence, respectively.

The following patterns frequently appear in the dependence between financial asset returns: (1) The upper and lower tails usually have different dependence patterns. (2) In both the upper and lower tails, dependence patterns are usually dynamic over time. Commonly used copulas have some limitations that make them unsuitable for modeling and comparing tail dependence among financial asset returns. Student-*t* copulas, for example, are widely used in the literature to quantify the dependence among stock returns due to their ability to model tail dependence. However, unlike full-range tail dependence copulas that will be discussed in further detail in what follows, Student-*t* copulas are always symmetric between their upper and lower tails; tail dependence exists not only in the upper and lower tails for capturing extreme positive or negative dependence, but also between extreme positive and negative returns. It is unlikely that we will have an extreme dependence on both returns of the same sign and returns of the opposite signs at the same time. Other commonly used copulas, such as the Gumbel, Clayton, and BB series copulas, can only model asymmetric dependence between the upper and lower tails.

When the dependence on the tails is not strong enough, λ_U or λ_L can be 0. Then, one can use $C(u, u) \approx u^{\kappa} \ell(u)$ as $u \to 0^+$ to model the tail dependence pattern, for which $\ell(u)$ is a slowly varying function, and the parameter $\kappa \ge 1$ is called the *tail order*, which is then used to describe the dependence strength in the tail; see Hua and Joe (2011) for details of tail orders. We only consider non-negative tail dependence, so the range of interested values of the tail order is $2 \ge \kappa \ge 1$.

When the lower tail dependence parameter $\lambda = 0$, the lower tail order $\kappa > 1$ can provide further information on the dependence strength of the lower tail. However, all the most popular copula models can only model one scenario: $\kappa = 1$ or $\kappa > 1$, but not both. To compare the strength of dependence in the upper and lower tails of intraday returns, we need a copula that has the following two critical features:

- Being able to capture the full-range tail dependence (i.e., the set of values of *κ* is dense in [1, 2]) in both the upper and lower tails. We refer to Definition 2.1 of Su and Hua (2017) for a formal definition of full-range tail dependence.
- 2. Being able to model both reflection symmetry and reflection asymmetry between the upper and lower tails. That is, the model can handle both cases: $\hat{C}(u,v) \equiv C(u,v)$ for any $(u,v) \in [0,1]^2$ and $\hat{C}(u,v) \neq C(u,v)$ with a positive probability for some $(u,v) \in [0,1]^2$. Note that the concepts of "reflection symmetry" and "permutation symmetry" are different and the latter means that $C(u,v) \equiv C(v,u)$ for any $(u,v) \in [0,1]^2$. We only consider copulas that have permutation symmetry that are mostly observed for dependence among returns of financial assets.

Constructing parametric copula families that have the above two common properties while also being computationally implementable has been a difficult task. So far, there are only two full-range tail dependence copulas: the GGEE and the PPPP copulas. Both are induced by X_i , i = 1, 2, which can be constructed by mixing four random variables of the following form:

$$X_i = (R_1/R_2) \cdot (Y_{i1}/Y_{i2}), \quad i = 1, 2,$$

where R_j 's and Y_{ij} 's are all independent. Since X_1 and X_2 share the same (R_1/R_2) and have independent (Y_{i1}/Y_{i2}) 's, the dependence between X_1 and X_2 becomes an interpolation of comonotonicty, the most positive dependence structure from the shared component (R_1/R_2) , and the independence structure. The names of the two copulas are based on what R_j 's and Y_{ij} 's are used: the PPPP copula is induced by such a mixture model of four Pareto random variables, and the GGEE copula is induced by two gamma and two exponential random variables. The GGEE copula was created in Hua (2017) and satisfies both the features 1 and 2 listed above. However, the current implementation in the R package CopulaOne¹ is quite slow in terms of computation speed. Later, the PPPP copula family was developed in Su and Hua (2017), which not only has the above features 1 and 2, but is also much faster computationally and is suitable for dealing with heavy computational tasks such as modeling dependence among high-frequency returns of financial assets.

The PPPP copula has four parameters α , a, β , b and the parameters enter the copula model as the shape parameters of the four Pareto random variables in the above stochastic representation. Since the shape parameter determines the tail heaviness of the Pareto random variable, an interaction between different levels of tail heaviness and the aforementioned interpolation between comonotonicity and independence leads to the flexible tail dependence structures. We refer to Assumption III of Su and Hua (2017) for mathematical details of the mixture model used to construct the PPPP copula. The normalized contour plots in Figure 2 of the PPPP copula have clearly demonstrated how flexible it can be. As a result, we will use the PPPP copula to model dependence patterns among financial asset returns.

The following summarizes some advantages of using full-range tail dependence copulas to model dependence patterns for financial asset returns: (1) It can capture dynamic dependence patterns in both the upper and lower tails. (2) It models the upper and lower tails separately with the corresponding parameter for the upper and lower tails, respectively, so that the degrees of dependence in the upper and lower tails can be modeled separately and compared directly without inherited distraction from the model itself. (3) It can account for both reflection symmetry and asymmetry, allowing the difference between the upper and lower tails to be driven solely by the data rather than the model constraints. (4) It can capture the full range of tail dependence, from tail independence to the strongest positive

tail dependence. (5) It can capture dynamic dependence with a single copula, eliminating the need for copula selection.



Figure 2. Normalized contour plots of the PPPP copula: The univariate marginals are transformed to the standard normal to help visualize the dependence patterns. The original PPPP copulas have four parameters: α , β , a, and b, and here we let a = b = 1 to further reduce the number of parameters so that the comparison between the upper and lower tails are only characterized by the remaining two parameters α and β for the lower and upper tails, respectively.

2.3. Unified Tail Dependence Measures

Tail dependence becomes very weak and close to independence as tail order κ approaches 2; see Hua and Joe (2011) for reference. When comparing values of tail orders close to 2, comparisons become less interesting and less informative. As a result, we can constrain the range of tail orders to the most interesting range [1, 2) by reparameterizing the parameters of the PPPP copula as follows: $\alpha = 1 + \frac{2}{\pi} \arctan(\alpha')$ and $\beta = 1 + \frac{2}{\pi} \arctan(\beta')$. Then, $\alpha', \beta' \in (-\infty, +\infty)$ and $\alpha, \beta \in (0, 2)$. We can achieve the ideal range of $\kappa \in [1, 2)$ according to Proposition 4.10 of Su and Hua (2017). It is especially useful when both the upper and lower tails have a very weak positive dependence and the full-range tail dependence copula is able to make comparing such weak positive dependence possible and informative.

We still have to deal with another issue after limiting the range of tail orders to [1,2). That is, we use the tail dependence parameters λ_L and λ_U to quantify the degree of tail dependence when the tail order $\kappa = 1$. If $\kappa > 1$, however, $\lambda_L = \lambda_U = 0$, and we must use κ_L and κ_U to measure the degree of tail dependence. With both κ and λ , comparisons between the cases of $\kappa = 1$ and $\kappa > 1$ become inconvenient; therefore, we need a unified quantity to allow such direct comparisons.

To address this issue, we propose to employ

$$\pi_L = \frac{2}{\pi} \arctan(1/\alpha) \text{ and } \pi_U = \frac{2}{\pi} \arctan(1/\beta), \quad 0 < \pi_L, \pi_U < 1, \tag{1}$$

where α and β are the corresponding parameters of the PPPP copula in order to quantify the strength of dependence in the lower and upper tails, respectively. In what follows, we refer to π_L and π_U as the *unified tail dependence measure* of the PPPP copula in the lower and upper tails, respectively.

The fully parameterized PPPP copula has four parameters: α , β , a, and b, and a comparison between the upper and lower tails is more sensible if a = b because a and b affect the strength of tail dependence as well and we do not want different values of a and b to distort the comparison between the upper and lower tails. We set a = b = 1 in the empirical study to focus on uncovering and comparing different tail dependence patterns, but not to aim to achieve a better-fitted model with potentially different values of a and b; letting a = b = 1 is just for convenience and serves our purpose well enough. On the one hand, the unified tail dependence measures of π_L and π_U change continuously in the two cases: $\kappa = 1$ and $1 < \kappa < 2$, and, on the other hand, the strength of dependence in the tails increases in the values of π_L and π_U for the lower and upper tails, respectively, regardless of whether $\kappa = 1$ or $1 < \kappa < 2$. When $0.5 \le \pi_L[\pi_U] < 1$, we have the usual lower [upper]

tail dependence with $\kappa = 1$, and when $0 < \pi_L[\pi_U] < 0.5$, we have intermediate lower [upper] tail dependence cases with $1 < \kappa < 2$. As a result, π_L and π_U are excellent metrics for capturing the degrees of dependence in the tails.

2.4. Multiple Components GARCH for Intraday Volatility

It is well known that temporal dependence exists in a single time series of financial prices or returns. To use copulas to assess the dependence patterns among different time series data, one needs to filter each time series data first, and then the copula is used to model residuals that are better than the original data to satisfy the identical and independent assumption required for fitting a copula.

As a result, we must first filter each time series data using appropriate models. Furthermore, when compared to volatility across days, intraday volatility of log returns has distinct characteristics, such as the fact that volatility is typically higher right after the market opens and before the market closes, while volatility is typically lower during lunchtime in the US market. We need a model that can account for both the seasonality of volatility within each trading day and the heterogeneity of volatility across trading days. In the following, we model log returns of each financial asset using a multiple components GARCH model.

Denote intraday log returns as $r_{t,i} := \ln(p_{t,i}) - \ln(p_{t,i-1})$, where $t \in \{1, 2, \dots, T\}$ is the day index with the total number of days being *T* and *i* is the index for time bars. Regular trading hours on US stock exchanges are from 9:30 a.m. to 4:00 p.m., New York time, totaling 6.5 regular trading hours. We use 30 s time bars, so $i = 2, 3, \dots, 780$ and 780 is the number of 30 s time bars in the regular trading hours of the market. We only keep the time bars where both data are available. Since we only consider stocks and ETFs that have high liquidity, missing values are not an issue in our study.

The multiple components GARCH has the following specification:

$$r_{t,i} = \mu_{t,i} + \epsilon_{t,i} \qquad \epsilon_{t,i} = (q_{t,i}\sigma_t s_i) z_{t,i}, \tag{2}$$

where $q_{t,i}$ is the stochastic component for the intraday volatility so that the innovation $z_{t,i}$ can be assumed as a standard distribution such as the standardized Student-*t* distribution; σ_t is the volatility for the specific day *t* and it can be modeled separately using some other models for daily volatility and then inserted into the model; and s_i represents the intraday seasonality for the *i*th time bar and it describes the average intraday seasonality effect over the days that we are modeling. We refer to Engle (2002) and Galanos (2023) for details of the standard GARCH model of multiple components and its implementation in the R package rugarch (Galanos 2023). Then, we use the PPPP copula to model the standardized residual $z_{t,i}$'s for different financial assets. For each stock, the standardized residuals for day *t* are converted into uniform scores $u_{t,i}$'s by their corresponding ranks within the day: $u_{t,i} = (\operatorname{rank}(z_{t,i}) - 0.5)/n_t$, where n_t is the total number of observations on day *t*. When we need to study a particular trading session, such as the 90 min period before the market closes, we retrieve the corresponding uniform scores are obtained on the basis of their daily ranks, even if we only use the data for a specific period within the day.

2.5. Regression Models with Full-Range Tail Dependence Copulas

To study how factors such as time of day and the Fama–French five factors affect intraday tail dependence, we perform regression analyses using the uniform scores of the standardized residuals from the aforementioned intraday marginal models as response variables, and the explanatory variables are linked via natural cubic splines to the upper and lower tail dependence parameters, respectively. Given observations *x* of the explanatory variables, we let the *x* enter the model via natural cubic splines as follows for both the upper and lower tail dependence parameters:

$$\alpha(\mathbf{x}) = 1 + \frac{2}{\pi} \arctan(\alpha'(\mathbf{x})), \quad \alpha'(\mathbf{x}) = \gamma_{\alpha 0} + \sum_{i=1}^{k} \gamma_{\alpha i} \cdot \operatorname{ns}_{i}(x_{i}) + \sum_{i=k+1}^{p} \gamma_{\alpha i} \cdot x_{i},$$

$$\beta(\mathbf{x}) = 1 + \frac{2}{\pi} \arctan(\beta'(\mathbf{x})), \quad \beta'(\mathbf{x}) = \gamma_{\beta 0} + \sum_{i=1}^{k} \gamma_{\beta i} \cdot \operatorname{ns}_{i}(x_{i}) + \sum_{i=k+1}^{p} \gamma_{\beta i} \cdot x_{i}$$

where x_1, \ldots, x_k are continuous explanatory variables with $ns_i(x_i)$ as the natural cubic splines for them and x_{k+1}, \ldots, x_p as the categorical explanatory variables. In this study, we only consider continuous explanatory variables, but categorical variables can also be used in the model. The parameters to be estimated are $\gamma = (\gamma_{\alpha}, \gamma_{\beta})$.

A maximum likelihood method can be used to obtain the estimates of these parameters, and the variance–covariance matrix of the estimated parameters can be derived from the approximated Hessian matrix. To assess the variability of the estimates, we can simulate regression parameters based on the estimates and the approximated variance–covariance matrix many times. Each time, we plot a line based on the simulated parameters (see the gray lines in Figures 5 and 7), and these simulated lines show the variability of the fitted regression lines.

3. Empirical Study

In this section, we present an empirical study for the ticker pairs in Table 1. These four pairs represent different categories of stock and ETF combinations and we can find similarities and differences in the dependence patterns between these categories. SPY is an ETF that tracks the S&P 500 index and it is one of the most popular equity ETFs. QQQ is an ETF that tracks the Nasdaq-100 index, with the largest holding at the time of the study being AAPL. Both MSFT and AAPL are dividend-paying, mega-cap growth stocks, while JPM and BAC are large-cap value stocks.

Table 1. Stock and ETF tickers studied.

Tickers	Categories	Tickers	Categories
SPY vs. QQQ	ETF vs. ETF	AAPL vs. MSFT	stock vs. stock (growth)
QQQ vs. AAPL	ETF vs. holdings	JPM vs. BAC	stock vs. stock (value)

The data set used for the empirical study is described in Section 3.1. Through the tickers of SPY and QQQ, the method details will be discussed in Section 3.2. Section 3.3 summarizes the findings on dynamic tail dependence patterns over time, and Section 3.4 presents the study of the effects of the Fama–French five factors on the unified tail dependence measures.

3.1. Data

The historical stock price data was obtained from Alpaca, a brokerage focused on algorithm trading. The 30 s time bars are constructed using the tick data for the year 2020, omitting odd-lot transactions and trades that occur during extended trading hours. The data for the whole year are used to study the effects of the Fama–French five factors, and Table 2 are summary statistics of the 30 s log returns.

A matched pair of market sell-off and market rally data is used to study the dynamic intraday tail dependence patterns over time. We use the performance of SPY as a proxy for the market performance. The market sell-off and market rally periods are selected from 20 February 2020 to 23 March 2020 and from 24 March 2020 to 24 April 2020, respectively. The period of market rally is matched to have the same number of trading days as those for the market crash period, although the market rally lasted longer than the chosen period. Figure 3 shows the daily closing prices during the chosen periods.

Ticker	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
SPY	$-1.0 imes 10^{-2}$	$-1.6 imes10^{-4}$	0.0	$3.5 imes 10^{-7}$	$1.6 imes 10^{-4}$	$1.7 imes 10^{-2}$
QQQ	$-1.1 imes10^{-2}$	$-2.1 imes10^{-4}$	0.0	$9.9 imes10^{-7}$	$2.2 imes 10^{-4}$	$1.5 imes 10^{-2}$
AAPL	$-2.4 imes10^{-2}$	$-3.0 imes10^{-4}$	0.0	$1.1 imes 10^{-6}$	$3.1 imes 10^{-4}$	$1.7 imes 10^{-2}$
MSFT	$-2.3 imes10^{-2}$	$-2.7 imes10^{-4}$	0.0	$8.2 imes10^{-7}$	$2.8 imes 10^{-4}$	$2.0 imes10^{-2}$
JPM	$-1.7 imes 10^{-2}$	$-3.0 imes10^{-4}$	0.0	$-1.0 imes10^{-6}$	$3.0 imes 10^{-4}$	$2.4 imes 10^{-2}$
BAC	$-1.9 imes 10^{-2}$	$-3.6 imes10^{-4}$	0.0	$6.6 imes10^{-7}$	$3.6 imes10^{-4}$	$2.8 imes 10^{-2}$

Table 2. Summary statistics of the 30 s log returns of the tickers under study during the year 2020. "1st Qu." and "3rd Qu." are the first and third quartiles, respectively.



Figure 3. Performance of SPY during the chosen periods of the market sell-off and the market rally in 2020.

The daily data for the Fama–French five factors of the US market were downloaded from Professor French's website.² The following factors are considered in the study: Mkt-RF (the excess return on the market), HML (High Minus Low: the average return on the two value portfolios minus the average return on the two growth portfolios), SMB (Small Minus Big: the average return on the nine small stock portfolios minus the average return on the nine big stock portfolios), CMA (Conservative Minus Aggressive: the average return on the two conservative investment portfolios minus the average return on the two aggressive investment portfolios), and RMW (Robust Minus Weak: the average return on the two robust operating profitability portfolios minus the average return on the two weak operating profitability portfolios). We refer to Professor French's website for detailed documents of the variables.

3.2. Intraday Marginal Analysis

First, we individually model each marginal time series data. To account for intraday volatility, we apply the standard GARCH model of multiple components described in Section 2.4, where the innovation is assumed to follow the standard Student-*t* distribution and GARCH (1,1) is used. ARMA(1,1) models the mean process of each marginal. Therefore, the overall model we use is the ARMA(1,1)-mcsGARCH(1,1, std) model. We utilized one month's worth of data to fit the model month by month. We have tried changing the orders of the time series models and the current one worked well, so we chose to use time series models that have the above orders for the marginal analysis of all the stocks.

The daily volatility σ_t in Equation (2) is modeled by an ARMA(1,1)-GARCH(1,1) model. A rolling window of size 1000 was used and the model was re-fitted every 5 days. Figure 4a shows the estimated daily volatility for 2020. It is clear that daily volatility spiked in March and then gradually calmed down in April.

The multiple components GARCH model accounts for intraday volatility patterns and Figure 4 shows relevant plots for SPY on 16 March 2020, which is one of the worst days in the history of the US stock market, where the S&P 500 index dropped about 12%, comparing the lowest in the day with the closing price of the previous day. Figure 4b shows the closing prices of SPY. Note that there was a 15 min halt at the market opening. Figure 4c depicts the 30 s log returns of SPY. Figure 4d shows the estimated seasonality of intraday volatility, which corresponds to s_i in Equation (2). Notice that after the market opens and just before it

closes, the volatility is notably greater. Figure 4e illustrates the estimated intraday stochastic component, which is $q_{t,i}$ in Equation (2). Figure 4f represents the standard residuals $z_{t,i}$ in Equation (2). For the residuals, we use the standard Student-*t* distribution.

We want to remark that, for intraday dependence modeling of multiple time series, it is very important to first use a model such as the multiple components GARCH model to account for the distinct marginal intraday features. Otherwise, the residuals are distorted by those intraday features, such as seasonality and stochastic components, and it would thus be misleading to use the corresponding intraday ranks of the residuals to calculate uniform scores for copula modeling. The idea can be understood by comparing Figure 4c,f, which clearly suggests that the standardized residuals are much more stationary than the log returns.



Figure 4. SPY's estimated daily volatility of the year 2020 and intraday plots for 16 March 2020. Note that there was a 15 min trading halt at the beginning of the regular trading hours due to the market crash. (a) Estimated daily volatility; (b) closing prices; (c) log returns; (d) intraday seasonality; (e) intraday stochastic components; and (f) intraday standardized residuals.

3.3. Dynamic Intraday Tail Dependence Patterns

After fitting the marginal time series, we can obtain the standardized residuals and then use them as pseudo-data to study the dependence between the corresponding time series. We are interested in knowing how the tail dependence changes over time during a trading day, under different market conditions and with different kinds of stocks. So, we use the regression model discussed in Section 2.5 and we use time units (in 5 min) since the opening of the market as the explanatory variable. The function ns from the R package splines was used to create the natural cubic splines, with two knots at one third and two thirds, respectively. The likelihood of the PPPP copula can be easily calculated using the R package CopulaOne, so we can use maximal likelihood estimators for the regression model. After the regression coefficients are obtained, intraday dependence patterns can then be inferred based on the estimated values of π_L and π_U .

In Figure 5, we plot the estimated unified tail dependence measures π_L and π_U for different kinds of stocks during the periods of a market sell-off and a market rally, respectively. The estimated unified tail dependence measures suggest the following:

- (a) Both the lower and upper tail dependence tend to peak towards the end of the regular trading hours of the day.
- (b) Both the lower and upper tail dependence tend to reach a plateau and/or slightly drop during the lunch session.

It is relatively easier to understand that during lunchtime, there are fewer trading activities, and thus, the dependence either stops increasing or starts to decrease mildly before the more active afternoon trading sessions. According to Koopman et al. (2018), the overall dependence among ten US banking equities in 2012 suggests that dependence levels tend to increase throughout the day. The authors attribute the dependence pattern to firm-specific information collected overnight, which results in a considerably smaller dependence at market opening. Another possible explanation is that there are more low-frequency traders or less sophisticated traders, such as retail traders, during morning sessions, thus associating trading activities with a greater number of market participants and resulting in a relative weakening of dependence. On the contrary, large institutional traders may trade many assets concurrently for things like fund rebalancing during the late afternoon sessions, causing dependence to increase at that time.

- 2. Dependence patterns attributed to market sentiments: For all the ticker pairs we have studied, market sentiments affect both the strength of overall dependence and the degree of asymmetry of dependence between the lower and upper tails:
 - (a) Overall dependence is stronger during a market sell-off by comparing the left two columns to the right two columns of Figure 5.
 - (b) The dependence is stronger in the upper tail than in the lower tail during a market sell-off, but it is not the case during a market rally.

Pattern (a) is consistent with what has already been studied in the literature and can be attributed to the financial risk contagious effect. Examples of such research work are Jondeau (2016), Rodriguez (2007), and Kato et al. (2022). To the best of our knowledge, pattern (b) has not been well studied in the literature and we are trying to discuss the potential reasons in what follows. When the market experiences a significant decline, the market sentiment is quite negative and fears penetrate the market. Therefore, there are many more sellers than buyers with market-moving power. As a result, the dependence between short-term (say 30 s) positive returns becomes stronger than that of short-term negative returns. The fewer buyers are probably the main forces driving the prices up in a short period of time. They could be intraday short sellers who tend to buy back those borrowed and sold stocks by the end of the day to avoid higher borrowing costs and overnight risks, and long term investors who may see market sell-offs as an opportunity to add to their equity positions. Interestingly, such an effect is not present during the market rally and it seems that the relative strength of dependence between the lower and upper tails also depends on other factors, such as time and stock types during a market rally.

3. Dependence patterns attributed to stock types: Stock types play a role in shaping intraday dependence patterns and general conclusions can not be drawn on dependence patterns for growth or value stocks. One needs to study them separately for any pair of stocks of interest.

The bottom row of Figure 5, for example, illustrates the tail dependence between two major bank stocks, which tends to be the highest immediately after the market opens and decreases throughout the day, with upward movements appearing after lunch. We hypothesize that the following factors may contribute to such a dependence pattern: (1) Unlike high-tech companies like AAPL and MSFT, banks like JPM and BAC have similar business and risk exposures; therefore, there is less overnight firm-specific

information that impacts only one bank but not the other. (2) Value equities, such as JPM and BAC, are traded in substantially lower volumes than popular large growth firms. During the study period, JPM's total adjusted traded value was just about 17% of AAPL's. Therefore, there would be distinct market players for growth and value stocks and it is reasonable to observe different intraday dependence patterns for different stock pairs.



Figure 5. Unified tail dependence measures change over time. From top to bottom, each row represents one pair of tickers and they are SPY vs. QQQ, QQQ vs. AAPL, AAPL vs. MSFT, and JPM vs. BAC, respectively. The two columns on the left are the plots for the market sell-off period from 20 February 2020 to 23 March 2020, and the two columns on the right are the plots for the market rally period from 24 March 2020 to 24 April 2020. The solid blue line is the estimated regression line, and the 200 gray lines are created based on simulated regression parameters to indicate the uncertainty of the estimated regression line.

3.4. The Fama–French Five Factors on Intraday Tail Dependence

To further explore factors that could affect intraday tail dependence patterns, we employ the regression model introduced in Section 2.5 again with the Fama–French five factors as explanatory variables. In what follows, we present the results for AAPL and MSFT during the last 90 min of each trading day in the year 2020. From the results obtained in Section 3.3, the intraday tail dependence tends to peak toward the end of the regular trading hours for all the stock/ETF pairs we have studied. Therefore, we use the data of the last 90 min as an illustration for our study.

Figure 6 shows the daily Fama–French five factors for 2020. For the natural cubic splines of the factors, we used two knots at one third and two thirds, respectively. We used the likelihood ratio test as the criterion for selecting the significant factors of the five factors. A forward selection approach was used to add one variable from the five factors each step; the one variable added each step was the one that makes the *p* value of the likelihood ratio test the smallest. We stopped adding more factors when no more factors could significantly improve the model in the sense that adding more factors will no longer make the *p* value less than 5%. Model selection is summarized in Table 3, where "df" is the degree of freedom and "chi-sq" is the test statistic of the likelihood ratio test. According to the table, the four factors of Mkt-RF, HML, SMB, and CMA significantly affect the intraday tail dependence and are included in the model. Polanski et al. (2021) uses daily returns data to examine the impact of Fama–French–Carhart factors on tail dependence among some stocks and finds that most of the dependence is accounted for by the Mkt-RF factor, while the other factors account for very little of the tail dependence. Our work considers high-frequency intraday data and also notices that the Mkt-RF plays the most important role in shaping the intraday tail dependence patterns. Moreover, our study also indicates that some other factors may also have a significant impact, although not as much as that of the Mkt-RF factor.



Figure 6. The daily Fama–French five factors for the year 2020.

In Figure 7, we plot the effects of the four significant factors on the lower tail dependence, the upper tail dependence, and the difference between the dependence of the upper and lower tails, respectively. The effect of each factor is presented while keeping the other factors as their corresponding median values. In what follows, we try to discuss the various dependence patterns and try to provide some potential reasons. However, a more comprehensive empirical study based on many other stocks from multiple years will be helpful in thoroughly understanding the mechanism of how these factors affect the tail dependence patterns. It is outside the scope of the paper and more research is required in this research direction.



Figure 7. Effects of the Fama–French five factors on the unified tail dependence measures. The solid blue line is the estimated regression line, and the 200 gray lines are created based on simulated regression parameters to indicate the uncertainty of the estimated regression line. It is clear that the variability of the estimated regression line for the factor Mkt-RF is smaller than that for the factor CMA.

It is interesting to note that both the upper and lower unified tail dependence measures reach the lowest when there is a positive excess return (Mkt-RF) of about 2% on the market, and the tail dependence increases in both the upper and lower tails when the excess returns move away from the "sweet spot" of the excess returns, regardless of the sign of the returns. A more negative excess return is associated with a larger gap between the unified upper and lower tail dependence measures. This is consistent with what we have discovered in Section 3.3, and the same potential reason for such a pattern discussed in Section 3.3 applies here.

 1.11×10^{-1}

The factor HML represents the difference between the returns of value and the growth stocks. Both the unified upper and lower tail dependence measures reach the minimum value when the absolute value of HML is small. A more negative HML is associated with a stronger lower tail dependence and a larger gap between the lower and upper tail dependence. The phenomenon might be attributed to large institutional investors selling large volumes of value stocks during a relatively shorter period of time, and mega-cap stocks like AAPL and MSFT being also treated as large value stocks at that moment.

The SMB factor represents the difference between the returns of small and large companies. A negative SMB is not significantly associated with changes in the unified tail dependence measures for both the upper and lower tails. It is reasonable considering that both AAPL and MSFT are large stocks. However, a more positive SMB is associated with stronger tail dependence in both the upper and lower tails of such large stocks.

The effects of the factor CMA, the difference between conservative and aggressive stocks, do not shape the tail dependence patterns as much as the above three factors, and, furthermore, the estimates have relatively greater variability indicated from the wider range of the gray lines (see the last row of Figure 7). Therefore, we restrain ourselves from overinterpreting it.

significance rever of 576.			
	df	chi-sq	<i>p</i> Value
Mkt-RF	6	459.86	$<1.0\times10^{-12}$
HML	6	141.94	$< 1.0 \times 10^{-12}$
SMB	6	56.90	$1.91 imes 10^{-10}$
СМА	6	19.40	3.54×10^{-3}

10.34

6

Table 3. Forward selection for the Fama–French five factors, where RMW is not significant at the significance level of 5%.

4. Concluding Remarks

RMW

By modeling the evolving intraday tail dependence patterns, it is possible to discover some common but hidden dependence patterns in the financial market. We recommend using the PPPP copula, a full-range tail dependence copula, to model standardized residuals for log return processes to capture dynamic tail dependence patterns between financial assets. To account for the unique characteristics of intraday volatility, a standard GARCH of multiple components model is required. To compare tail dependence at various levels, a model-based unified tail dependence measure is suggested. Instead of using different measures in different situations, this metric enables direct comparisons of tail dependence strength in various situations. Through the empirical study, we have uncovered several interesting intraday tail dependence patterns for popular US stocks and ETFs over time via explanatory variables such as the Fama–French five factors. We have attempted to explain the reasons for these dependence patterns, but more empirical research is required to justify them. The proposed method and the unified tail dependence measures can also be applied to other financial assets.

The study of intraday tail dependence in financial markets provides more aspects of the market to both investors and policymakers. It can help them better understand the sentiments of the intraday market and the behavior of market participants. It offers insights into extreme events, which are often the most challenging and important for risk assessment and policy decisions. The current study is limited to the year 2020 and the selected ticker symbols. However, the proposed method can be applied to many other research questions. For example, a significant asymmetry between the upper and lower tail dependence could be helpful in detecting manipulative market behavior and the footprints of influential market players. Future studies can be carried out to examine a large number of stocks over multiple years. Such a study would be helpful to better understand the dynamic tail dependence patterns in different market environments and for different categories of financial assets and thus help prevent and manage large intraday financial risks.

Funding: This research received no external funding.

Data Availability Statement: Data sharing is not applicable to this article.

Conflicts of Interest: The author declares no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AAPL	The ticker symbol for Apple Inc.
ARMA	Auto Regressive Moving Average
BAC	The ticker symbol for Bank of America Corporation
BB	A series of 2-parameter copulas in Section 5.2 of Joe (1997)
CDF	Cumulative Distribution Function
СМА	Fama–French five factors: Conservative Minus Aggressive
ETF	Exchange-traded Fund
GARCH	Generalized Auto Regressive Conditional Heteroskedasticity
GGEE	Full-range tail dependence copula by 2 Gamma & 2 Exponential random variables
HML	Fama–French five factors: High Minus Low
PM	The ticker symbol for JPMorgan Chase & Co.
Mkt-RF	Fama–French five factors: The excess return on the market
MSFT	The ticker symbol for Microsoft Corporation
PPPP	Full-range tail dependence copula by 4 Pareto random variables
QQQ	A ETF that tracks the performance of the NASDAQ-100 Index
RMW	Fama–French five factors: Robust Minus Weak
SMB	Fama–French five factors: Small Minus Big
SPY	A ETF that tracks the performance of the S&P 500 Index
US	United States

Notes

- ¹ https://github.com/larryleihua/CopulaOne (accessed on 4 November 2023).
- ² https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_5_factors_2x3.html (accessed on 4 November 2023).

References

Buccheri, Giuseppe, Giacomo Bormetti, Fulvio Corsi, and Fabrizio Lillo. 2021. A score-driven conditional correlation model for noisy and asynchronous data: An application to high-frequency covariance dynamics. *Journal of Business Economic Statistics* 39: 920–36. [CrossRef]

Carvalho, Marcela de Marillac, Luiz Otávio de Oliveira Pala, Gabriel Rodrigo Gomes Pessanha, and Thelma Sáfadi. 2021. Asymmetric dependence of intraday frequency components in the brazilian stock market. *SN Business Economics* 1: 1–18. [CrossRef] [PubMed]

- De Luca, Giovanni, and Paola Zuccolotto. 2017. Dynamic tail dependence clustering of financial time series. *Statistical Papers* 58: 641–57. [CrossRef]
- Dißmann, Jeffrey, Eike Christian Brechmann, Claudia Czado, and Dorota Kurowicka. 2013. Selecting and estimating regular vine copulae and application to financial returns. *Computational Statistics Data Analysis* 59: 52–69. [CrossRef]

Engle, Robert. 2002. New frontiers for arch models. Journal of Applied Econometrics 17: 425–46. [CrossRef]

Fama, Eugene F., and Kenneth R. French. 2015. A five-factor asset pricing model. Journal of Financial Economics 116: 1–22. [CrossRef]

Galanos, Alexios. 2023. Rugarch: Univariate Garch Models. R Package Version 1.5-1. Available online: https://cran.r-project.org/ web/packages/rugarch (accessed on 4 November 2023).

- Hua, Lei. 2017. On a bivariate copula with both upper and lower full-range tail dependence. *Insurance: Mathematics and Economics* 73: 94–104. [CrossRef]
- Hua, Lei, and Harry Joe. 2011. Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis* 102: 1454–71. [CrossRef]

Joe, Harry. 1997. Multivariate Models and Dependence Concepts. London: Chapman & Hall. [CrossRef]

Jondeau, Eric. 2016. Asymmetry in tail dependence in equity portfolios. *Computational Statistics Data Analysis* 100: 351–68. [CrossRef] Kato, Shogo, Toshinao Yoshiba, and Shinto Eguchi. 2022. Copula-based measures of asymmetry between the lower and upper tail probabilities. *Statistical Papers* 63: 1907–29. [CrossRef]

- Kim, Jong-Min, and Sun Young Hwang. 2021. Functional ARCH directional dependence via copula for intraday volatility from high-frequency financial time series. *Applied Economics* 53: 506–20. [CrossRef]
- Koopman, Siem Jan, Rutger Lit, André Lucas, and Anne Opschoor. 2018. Dynamic discrete copula models for high-frequency stock price changes. *Journal of Applied Econometrics* 33: 966–85. [CrossRef]
- Krupskii, Pavel, and Harry Joe. 2020. Flexible copula models with dynamic dependence and application to financial data. *Econometrics* and Statistics 16: 148–67. [CrossRef]
- Patton, Andrew J. 2009. Copula-based models for financial time series. In *Handbook of Financial Time Series*. Berlin/Heidelberg: Springer, pp. 767–85. [CrossRef]

Patton, Andrew J. 2012. A review of copula models for economic time series. Journal of Multivariate Analysis 110: 4–18. [CrossRef]

Polanski, Arnold, Evarist Stoja, and Ching-Wai Chiu. 2021. Tail risk interdependence. *International Journal of Finance Economics* 26: 5499–11. [CrossRef]

Rodriguez, Juan Carlos. 2007. Measuring financial contagion: A copula approach. Journal of Empirical Finance 14: 401–23. [CrossRef]

- Sklar, Abe. 1959. Fonctions de répartition à *n* dimensions et leurs marges. *Publications de l'Institut de statistique de l'Universite' de Paris* 8: 229–31.
- Su, Jianxi, and Lei Hua. 2017. A general approach to full-range tail dependence copulas. *Insurance: Mathematics and Economics* 77: 49–64. [CrossRef]
- Weiß, Gregor N. F., and Hendrik Supper. 2013. Forecasting liquidity-adjusted intraday value-at-risk with vine copulas. *Journal of Banking Finance* 37: 3334–50. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.