

Copula and Regression Based Analysis of the S&P 500 and VIX Dependence Structure

EC4305 Applied Econometrics

Jeron Tan Kang, Cai Mun Jia, Kellianne Ng, Khalisah Binte Shari, Tammy Alexandra Wong

December 2025

Contents

1	Introduction	1
2	Related Literature for Distribution and Dependence Analysis	1
3	Data & Target Variables	2
3.1	Data Overview	2
3.2	Preprocessing	4
4	Summary Statistics	4
4.1	Methodology	4
4.2	Results	4
5	Distribution Tests	6
6	Model Estimation	6
7	Moment Generating Function Estimates	8
8	Univariate Empirical Distributions	8
9	Probability Integral Transformed Data	9
10	Copula Estimation	10
10.1	From Linear Correlation to Copulas	11
10.2	Tail Dependence Functionals	11
10.3	Parametric Copula Benchmarks	12
10.3.1	Elliptical Copulas: Gaussian, Student- t , and Skew- t	12
10.3.2	Archimedean Copulas and Rotations	16
10.4	Exponential Series Estimator Copula	24
10.4.1	Model Specification	24
10.4.2	Estimation Procedure	25
10.4.3	Estimated Copula Surface and Interpretation	25

10.5 Comparison of Copula Models	27
11 Limitations for Distribution and Dependence Analysis	31
12 Conclusion for Distribution and Dependence Analysis	31
13 Regression Analysis	32
13.1 Linear & Non-Linear models	33
13.1.1 S&P 500 returns on VIX returns	33
13.1.2 S&P 500 returns on squared of VIX returns	33
13.2 Interaction	34
13.3 Autoregressive Model	36
13.4 Autoregressive Distributed Lag (ADL) Model	37
13.5 Instrumental Variable Estimation	39
13.5.1 Economic Policy Uncertainty Index	39
13.5.2 Lag of rvix	41
14 Conclusion for Regression Analysis	42

1 Introduction

The S&P 500 Index, a market capitalization weighted index of 500 of the largest publicly traded U.S. companies, is universally recognized as the premier benchmark for the health and performance of the U.S. equity market. Its counterpart, the CBOE Volatility Index (VIX), is widely known as the “investor’s fear gauge” because it reflects the market’s expectation of near-term volatility. The VIX is a forward-looking measure, calculating the 30-day implied volatility extracted from a broad range of S&P 500 index call and put options. Elevated VIX levels signal greater expected price swings in the S&P 500 and are intrinsically linked to periods of heightened uncertainty or risk aversion, establishing a pervasive, though often non-linear, inverse relationship between the two indices. This role as a barometer of market stress was dramatically highlighted during the Global Financial Crisis of 2008, when the VIX reached a then-record closing high of 80.86, signaling the unprecedented turmoil faced by global stock markets¹.

In this study, we aim to move beyond simple correlation to fully understand the dynamic relationship between S&P 500 returns (rsp) and VIX returns ($rvix$). Modeling this relationship is empirically challenging due to the inherent properties of the data. Both return series are highly non-normal, displaying significant skewness and excess kurtosis (as shown in Section 4). This non-Gaussian nature, combined with the critical observation that the dependence structure is inherently nonlinear and state-dependent—with co-movements intensifying during market stress (specifically, the negative dependency is stronger in extreme bearish markets than in extreme bullish markets)—fundamentally limits the effectiveness of traditional, linear correlation measures and Gaussian-based models in capturing joint crash risk. Furthermore, empirical phenomena such as volatility clustering induce serial dependence and conditional heteroskedasticity, undermining standard inference.

This paper examines the dependence between S&P500 and VIX returns in two complementary steps. We first attempt to model the joint behaviour of S&P 500 and VIX returns using a skew- t distribution, motivated by the finding in Section 6 that the skew- t distribution provides the best marginal fit among the candidate models. However, we find that even the skew- t framework cannot fully capture the pronounced asymmetry and heavy tails observed in VIX returns. Section 8 therefore employs both parametric and nonparametric copula approaches to model the joint distribution of S&P500 and VIX returns, allowing for flexible, potentially nonlinear dependence. Section 9 then analyses the relationship through a series of regression models, which provide an interpretable, conditional perspective on how S&P 500 returns respond to movements in VIX returns. Finally, Section 10 uses the fitted copulas to compute and compare tail functionals—including conditional tail dependence indices and joint exceedance probabilities—so that each model can be evaluated on its ability to reproduce “crash–spike” co-movements in a common, margin-free scale. As a whole, these analyses offer a more comprehensive view of the dependence structure between the two return series.

2 Related Literature for Distribution and Dependence Analysis

A large literature documents a strong and state dependent negative dependence between equity returns and implied volatility, often emphasising that linear correlation is an incomplete summary

¹The VIX closed at 80.86 on 20 November 2008.

because it cannot separate marginal non normality from dependence and cannot describe tail co movements. In the context of the S&P 500 and the VIX, this motivates copula based and other nonparametric approaches that explicitly target nonlinear and tail features of the joint distribution.

(Sun and Wu, 2009) study daily S&P 500 returns and VIX returns using a model free *nonparametric copula* framework and introduce a *conditional dependence index* that varies across segments of the market-return distribution. They report three core findings: (i) a strong negative dependence overall, (ii) pronounced *extreme tail dependence* consistent with crash–spike episodes, and (iii) clear *state dependence*, with the negative dependence being stronger in extreme bearish markets than in extreme bullish markets, and weakening as returns move toward the center of the distribution (quiet markets). They further use these insights to motivate simple predictive specifications for VIX changes that incorporate tail indicators, highlighting the economic role of VIX as a “fear gauge” that reacts disproportionately during equity stress.

Complementing copula based evidence, (Allen et al., 2013) adopt nonparametric density estimation and *entropy-based* metrics (including mutual information) to examine the S&P 500–VIX relationship across subperiods. Their results suggest that the marginal and joint distributional features of returns change shape across time (notably around crisis periods), and formal entropy based nonparametric tests reject distributional equivalence and symmetry in relevant subsamples. They conclude that purely parametric techniques can miss important complexities and emphasise the practical difficulty of calibrating hedges using VIX derivatives when dependence is time varying and tail sensitive.

Recent copula model selection exercises also point toward *rotated* families as a parsimonious way to encode negative tail dependence. For example, (Rašiová and Árendáš, 2022) fit elliptical and Archimedean copulas after filtering serial dependence via ARMA–GARCH residuals, and find that a 270° rotated Gumbel copula can be preferred by cross validation for capturing intensified dependence in extreme states (high volatility and low equity sector returns). Although their application is sector-specific, their approach reinforces the broader motivation for considering rotated copulas when modelling negative “crash–spike” type dependence.

Finally, our nonparametric benchmark uses the *Exponential Series Estimator* (ESE) of (Wu, 2007), which models a multivariate density by exponentiating a finite series expansion of basis functions and selects tuning parameters using information criteria. A key practical advantage highlighted in (Wu, 2007) is that the ESE can substantially outperform kernel estimators in finite samples, in part because kernel methods suffer boundary bias on bounded supports. The paper also emphasises that once a smooth joint density and marginals are available, dependence measures such as mutual information become straightforward to compute and inherit desirable invariance properties under monotone transformations. In our setting, implementing an ESE copula offers a flexible way to approximate the copula density on $[0, 1]^2$ without imposing the symmetry or tail shape restrictions inherent in low dimensional parametric families, while still permitting direct comparison via common tail and joint extremes diagnostics.

3 Data & Target Variables

3.1 Data Overview

The dataset used in this study consists of daily return series for the S&P 500 and the VIX, calculated using the log-difference transformation detailed in the preprocessing step. The final

returns sample obtained spans from 3 January 1992 to 27 October 2025, comprising 8,515 daily observations that correspond to U.S. trading days. This long time horizon captures a wide range of market conditions—including major volatility events such as the dot-com crash in 2000, the 2008 global financial crisis, the COVID-19 pandemic in 2019, and post-pandemic regime shifts—making it well-suited for studying the dynamic and potentially asymmetric relationship between equity returns and implied volatility. By focusing on returns rather than price levels, the analysis ensures stationarity and facilitates interpretation of short-run dynamics.

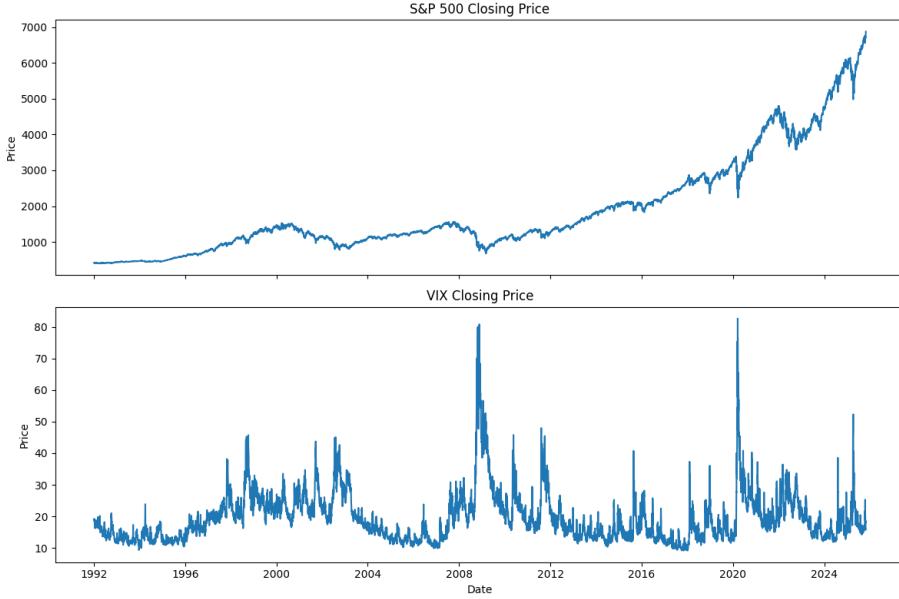


Figure 1: S&P500 and VIX over time

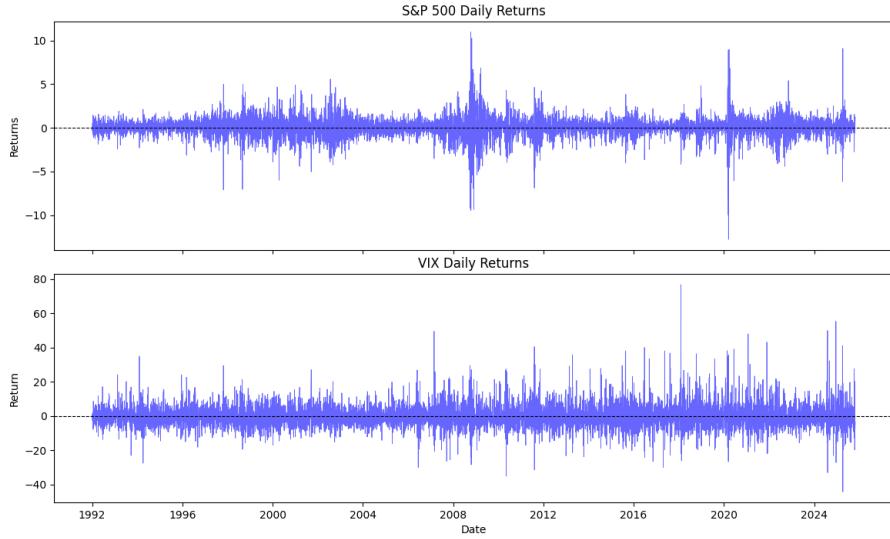


Figure 2: S&P500 returns and VIX returns over time

Table 1: Summary statistics for S&P 500 Index Level and VIX.

	Min	Max	Mean	Median	Std. Dev.	5th Pctl.	95th Pctl.
S&P500	394.50	6875.16	1879.06	1331.02	1396.58	450.09	4995.77
VIX	9.14	82.69	19.39	17.42	7.93	11.35	33.39

Table 2: Summary statistics for daily S&P 500 returns (*rsp*) and VIX returns (*rvix*).

	Min	Max	Mean	Median	Std. Dev.	5th Pctl.	95th Pctl.
<i>rsp</i>	-12.7652	10.9572	0.0329	0.0607	1.1504	-1.7615	1.6351
<i>rvix</i>	-44.2449	76.8245	-0.0021	-0.4303	6.8030	-9.7658	10.9788

3.2 Preprocessing

We first obtained daily data on the S&P 500 index level and the CBOE Volatility Index (VIX) from Yahoo Finance via the `yfinance` Python package. For both series, we retain only the daily closing prices on trading days, ensuring that observations are aligned across markets. The sample spans from 2 January 1992 to 27 October 2025, yielding 8,516 matched trading-day observations for each index.

Based on the daily closing prices for the S&P 500 (P_t) and VIX (VIX_t), we calculated the daily returns for both variables using the following transformation:

$$rsp_t = 100 \times \ln\left(\frac{P_t}{P_{t-1}}\right) \quad \text{and} \quad rvix_t = 100 \times \ln\left(\frac{VIX_t}{VIX_{t-1}}\right).$$

4 Summary Statistics

4.1 Methodology

Daily returns of the S&P 500 index (*rsp*) and VIX (*rvix*) were analyzed to summarize their distributional characteristics. We computed the mean, variance, skewness, kurtosis, and the correlation between the two using the `scipy.stats` package.

4.2 Results

Table 3: Summary Statistics

	Mean	Variance	Skewness	Kurtosis
<i>rsp</i>	0.032906	1.323458	-0.375141	11.061215
<i>rvix</i>	-0.002142	46.280273	0.956732	6.825242
Pearson Correlation (rsp, rvix): -0.722163				

Mean. The S&P 500 exhibits a small positive average daily return (≈ 0.03), consistent with the presence of an equity risk premium over the long run. The VIX, which reflects implied volatility, shows a slightly negative average return (≈ -0.002), indicating gradual mean reversion or a mild downward drift in volatility following spikes.

Variance. The variance of VIX returns (≈ 46.3) is substantially higher than that of S&P 500 returns (≈ 1.3). The VIX measures the market's expectation of future S&P 500 volatility, acting as a "fear gauge" that is extremely sensitive to investor sentiment. This leads to dramatically amplified movements, particularly sharp spikes during market downturns, resulting in a much higher variance in returns compared to the S&P 500's more gradual price changes.

Skewness. S&P 500 returns are mildly negatively skewed (≈ -0.38), reflecting a higher probability of large negative moves than large positive ones (crash risk). In contrast, VIX returns are positively skewed (≈ 0.96): the index occasionally experiences large upward jumps when volatility surges during equity sell offs.

Kurtosis. Both series display excess kurtosis, because extreme market movements occur more frequently than a normal distribution would predict. The S&P 500's excess kurtosis (≈ 11.1) is much higher than VIX (≈ 6.8), consistent with VIX's mean reverting tendency.

Correlation. The Pearson correlation between S&P 500 and VIX returns is strongly negative (≈ -0.72), indicating that when equities fall, implied volatility typically rises. The relation is asymmetric: the negative co movement strengthens during market downturns, as they are often accompanied by a rapid increase in investor anxiety which pushes VIX even higher.

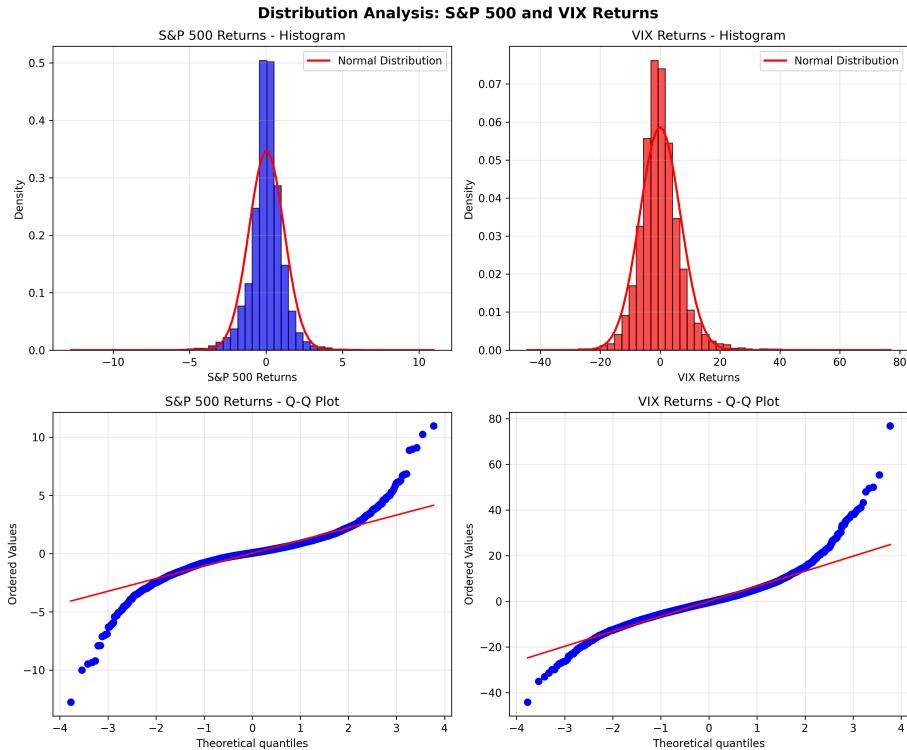


Figure 3: Histogram & Q-Q Plots for S&P 500 and VIX

The histograms for both the S&P 500 and VIX returns show distributions centered around a near zero mean, but with distinct characteristics. The S&P 500 histogram is more compact along the x-axis while the VIX returns is substantially wider, visually confirming its significantly higher variance. Both plots exhibit a sharp central peak that is much higher than the overlaid normal distribution curve, indicating high kurtosis.

In the Q-Q plots, both series display systematic tail deviations, indicating heavy tails and non normality. For the S&P 500, the lower left tail bends away from the line more than the upper right tail, consistent with negative skewness and high excess kurtosis. For VIX, the upper right tail shows the strongest deviation, consistent with positive skewness.

5 Distribution Tests

To formally assess whether S&P 500 and VIX return distributions follow a Normal distribution, two complementary normality tests were applied: the **Jarque–Bera (JB)** and **Anderson–Darling (AD)** tests. The JB test examines departures from Normality via skewness and kurtosis, while the AD test focuses on tail behavior in the empirical distribution.

Hypotheses

For both return series r_t^{SP} (S&P 500) and r_t^{VIX} :

H_0 : $r_t^{\text{SP}}, r_t^{\text{VIX}}$ are i.i.d. and normally distributed with mean μ and variance σ^2 ,

H_1 : The returns do not follow a normal distribution.

Test Statistics

Jarque–Bera (JB) test. For sample size n , sample skewness S , and sample kurtosis K (not excess):

$$JB = \frac{n}{6} \left(S^2 + \frac{(K - 3)^2}{4} \right),$$

where under H_0 , $JB \sim \chi^2_{(2)}$. Large JB indicates departures from Normality.

Anderson–Darling (AD) test.

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^n (2i - 1) [\ln F(z_i) + \ln(1 - F(z_{n+1-i}))],$$

where z_i are ordered standardized observations and F is the CDF of the standard Normal. Larger A^2 implies stronger evidence against Normality, especially in the tails.

Results

- **S&P 500:** $JB = 43,552.10, p < 0.01$; $AD = 157.67, p < 0.01$.
⇒ Reject H_0 ; distribution is non Normal.
- **VIX:** $JB = 17,803.34, p < 0.01$; $AD = 83.16, p < 0.01$.
⇒ Reject H_0 ; distribution is non Normal.

Both series exhibit non Normal, heavy tailed, and asymmetric distributions. These findings motivate the use of alternative heavy tailed or skewed models (e.g., Student- t , skew- t , Laplace) for modeling the distributions in subsequent analysis.

6 Model Estimation

Methodology

To capture the non Normal behavior observed in the S&P 500 and VIX return distributions, four candidate distributions were fitted via maximum likelihood estimation (MLE):

1. Gaussian (Normal) — baseline symmetric, thin-tailed model.
2. Student- t — allows for heavy tails through the degrees-of-freedom parameter ν .
3. Skew- t — adds a skewness parameter λ , accommodating both asymmetry and tail thickness.
4. Laplace — symmetric, with sharper peak and heavier tails than Gaussian.

For each model, parameters $\hat{\theta}$ maximize the log likelihood:

$$\ell(\hat{\theta}) = \sum_{t=1}^n \ln f(r_t | \hat{\theta}),$$

where $f(\cdot)$ denotes the probability density function. Model fit is compared using

$$\text{AIC} = 2k - 2\ell(\hat{\theta}), \quad \text{BIC} = k \ln(n) - 2\ell(\hat{\theta}),$$

where k is the number of parameters and n the sample size. The lowest AIC/BIC indicates the preferred specification.

Results

Table 4: Model comparison for S&P 500 returns

Distribution	Parameters	Tail Parameter	Log-Likelihood	AIC	BIC
Gaussian	2	—	-13,274.92	26,553.83	26,567.93
Student- t	3	df = 2.71	-12,049.89	24,105.78	24,126.93
Skew- t	4	a = 1.31	-12,040.63	24,089.25	24,117.45
Laplace	2	—	-12,088.19	24,180.39	24,194.49

Table 5: Model comparison for VIX returns

Distribution	Parameters	Tail Parameter	Log-Likelihood	AIC	BIC
Gaussian	2	—	-28,048.06	56,820.13	56,834.23
Student- t	3	df = 3.82	-27,770.42	55,446.84	55,467.99
Skew- t	4	a = 2.26	-27,680.26	55,368.55	55,396.75
Laplace	2	—	-27,794.48	55,592.96	55,607.06

Across both return series, the Skew- t distribution provides the best fit (lowest AIC/BIC). The Gaussian model performs worst due to its thin tails and symmetry, while the Student- t improves fit by capturing fat tails but misses asymmetry. Laplace captures kurtosis but remains symmetric.

Looking at the tail parameters, we note that both Student- t and Skew- t exhibit heavy tails. From our analysis, we observe that the VIX returns are more tail heavy than S&P 500. This discrepancy is linked to the VIX's function as an aggregate measure of implied volatility, heavily influenced by the vega of the underlying options portfolio. During a market downturn (a tail event for the S&P 500), investor demand for portfolio protection surges, driving significant hedging flows that rapidly increase option prices. This reflexive increase in implied volatility amplifies the VIX's upward movement, resulting in more pronounced positive return spikes compared to the S&P 500's own tail events.

7 Moment Generating Function Estimates

Based on the model selection results in the previous section, we use the Skew- t distribution to model the marginal distributions of the S&P 500 returns, r_t^{SP} , and the VIX returns, r_t^{VIX} .

To further assess whether the Skew- t model adequately captures the tail behaviour of each series, we examine their moment generating functions (MGFs). The rate at which the MGF increases reflects how heavy the tails of the distribution are – the higher the rate, the heavier the tails. This provides a useful check to see if the fitted Skew- t distribution matches the extreme tail behaviour in the observed data.

The MGF of the Skew- t distribution does not have a closed form, so we compute the empirical MGF instead. For a given value of t , the empirical MGFs are approximated using the observed data and the following formulas:

$$M_r^{\text{SP}}(t) = \mathbb{E} \left[e^{tr_t^{\text{SP}}} \right], \quad M_r^{\text{VIX}}(t) = \mathbb{E} \left[e^{tr_t^{\text{VIX}}} \right],$$

Results

Series	$M(-1)$	$M(0)$	$M(1)$
S&P 500	51.31	1.00	15.51
VIX	1.93×10^{15}	1.00	2.72×10^{29}

Table 6: Empirical MGF Estimates for S&P 500 and VIX Returns

The empirical MGF values highlight a clear difference in the tail behaviour of the two series.

S&P 500 Returns. The MGF values at $t = -1$ and $t = 1$ are relatively small, indicating lighter tails and a lower likelihood of extreme price movements. This is consistent with the typical behaviour of equity index returns, which tend to exhibit moderate rather than extreme tail risk.

VIX Returns. The MGF values at $t = -1$ and $t = 1$ are extremely large, indicating very heavy tails and a much higher frequency of extreme movements. This pattern reflects the tendency of volatility indices to experience sharp spikes during periods of market stress, a feature that aligns with the jump-like and asymmetric dynamics commonly observed in volatility data.

Overall, the MGF results reinforce the distinct risk profiles of the two series and highlight why modelling VIX returns often requires substantially heavier tailed distributions than those used for equity, like S&P 500, returns.

8 Univariate Empirical Distributions

To evaluate the fit of the Skew- t model, we compare the empirical distributions of the S&P 500 and VIX returns with their fitted Skew- t probability density functions (PDFs). We generate histograms for each series and overlay the corresponding fitted PDFs.

Results

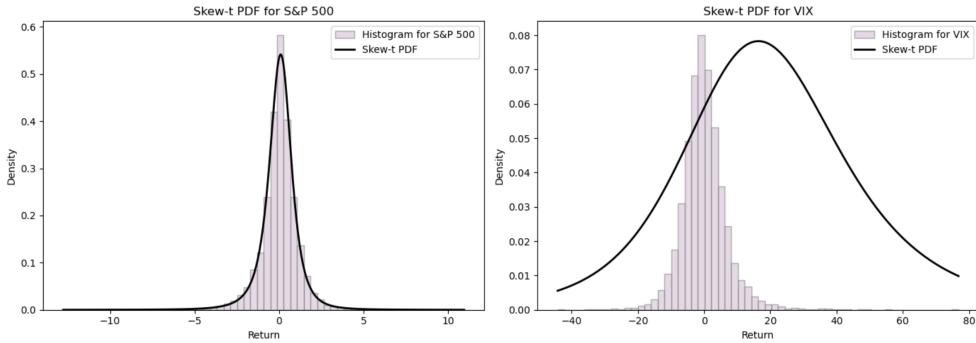


Figure 4: Histograms of S&P 500 and VIX returns with fitted Skew-*t* density curves.

S&P 500 Returns. The empirical distribution returns exhibits mild asymmetry and moderately heavy tails. The fitted Skew-*t* PDF captures these features reasonably well, indicating good alignment with the empirical distribution. The empirical MGF results also indicate moderate tail growth, which is consistent with the level of tail heaviness that the Skew-*t* distribution is designed to model. Taken together, these findings support the suitability of the use of Skew-*t* to model the S&P 500 returns.

VIX Returns. The empirical distribution exhibits strong asymmetry and extremely heavy tails. The fitted Skew-*t* PDF does not capture these features well, indicating a clear misalignment with the empirical distribution. The empirical MGF results also show very rapid tail growth, which exceeds the level of tail behaviour that the Skew-*t* distribution is designed to model. Taken together, these findings suggest that the Skew-*t* specification is not suitable for modelling the distributional characteristics of VIX returns.

9 Probability Integral Transformed Data

Integral Transformation

To prepare for the copula modelling, we apply probability integral transformation (PIT) on the S&P 500 and VIX returns using their estimated marginal cumulative distribution functions (CDFs). This transformation standardises the two series by removing the effects of their marginal distributions, so that any remaining structure in the data reflects only their dependence. Each return is transformed into a uniform variable on $[0, 1]$:

$$u_{SP} = F_{SP}(r_t^{SP}), \quad u_{VIX} = F_{VIX}(r_t^{VIX}).$$

Scatter Plot of Transformed Data

We then prepare a scatter plot of the transformed data points (u_{SP}, u_{VIX}) . This allows us to visually examine the dependence between S&P 500 and VIX returns, such as whether it is symmetric, asymmetric, or concentrated in the tails.

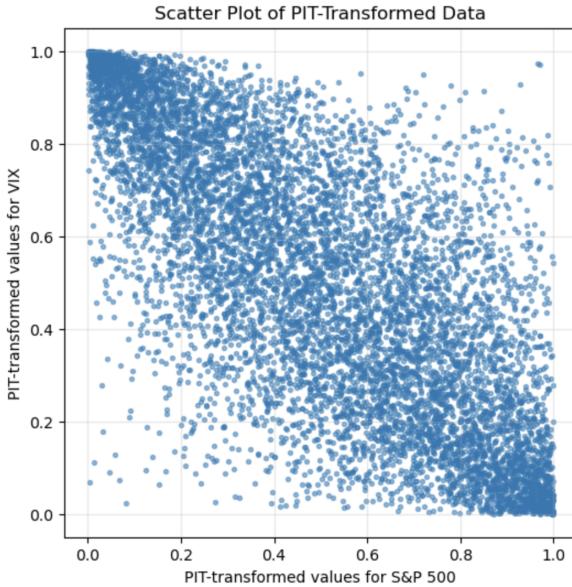


Figure 5: Scatter plot of PIT transformed S&P 500 and VIX returns.

Results

The scatter plot reveals a clear negative dependence between the two markets. High PIT values for S&P 500 tend to coincide with low PIT values for VIX. Conversely, low PIT values for the S&P 500 tend to coincide with high PIT values for VIX. This pattern reflects the well known inverse relationship between equity markets and volatility – when equity prices decline, volatility typically rises sharply, and when equity prices increase, volatility tends to fall.

The shape of the plot also suggests that the dependence may be stronger in the upper left and lower right regions, indicating possible asymmetry or tail dependence. Such features are important for modelling because linear correlation alone cannot capture them. These PIT transformed observations therefore provide a clearer view of the joint behaviour of the two series, independent of their marginal distributions, and form the basis for the copula estimation in the following section, where we model this dependence structure formally.

10 Copula Estimation

In this section we model the joint dependence between S&P 500 returns, r_t^{SP} , and VIX returns, r_t^{VIX} , using copulas. Unconditional summary statistics for both series indicate strong deviations from normality, including negative skewness and excess kurtosis for equity returns and pronounced right skewness for VIX returns, together with clear volatility clustering. These features imply that any analysis based solely on a single linear correlation coefficient is unlikely to provide an adequate description of the dependence structure, especially in the tails.

The copula framework separates marginal behaviour from dependence. We first map both series to a common probability scale using their empirical distribution functions and construct rank based pseudo observations. These pseudo observations define the empirical copula, which provides a purely nonparametric benchmark for the dependence structure. We then fit several standard parametric copulas (Gaussian, Student- t , and selected Archimedean families and rotations) and use them as parsimonious benchmarks for tail dependence and asymmetry. Finally, we implement

the Exponential Series Estimator (ESE), a flexible nonparametric copula model, and compare its implied tail behaviour to both the empirical copula and the parametric benchmarks (Wu, 2007).

10.1 From Linear Correlation to Copulas

A common starting point is the Pearson correlation coefficient between r_t^{SP} and r_t^{VIX} , which measures linear comovements around the mean. In the present setting this approach is unsatisfactory for at least three reasons.

First, the marginal distributions of returns are markedly non normal. The correlation coefficient is sensitive to heavy tails and skewness, which means that it mixes marginal distributional features with dependence and is not directly comparable across different marginal specifications. Second, correlation is not invariant to strictly increasing transformations. Transforming VIX returns by a log or variance stabilising map may alter the reported correlation even if the underlying economic dependence is unchanged. Third, correlation is fundamentally a measure of linear association and is not designed to capture nonlinear features such as tail dependence, where large negative moves in the S&P 500 are associated with large positive spikes in the VIX.

To address these limitations we adopt the copula representation from Sklar's theorem. Let F_{SP} and F_{VIX} denote the marginal cumulative distribution functions (CDFs) of r_t^{SP} and r_t^{VIX} , and let $F_{\text{SP},\text{VIX}}$ denote their joint CDF. Sklar's theorem states that there exists a copula function $C : [0, 1]^2 \rightarrow [0, 1]$ such that

$$F_{\text{SP},\text{VIX}}(x, y) = C(F_{\text{SP}}(x), F_{\text{VIX}}(y))$$

for all $(x, y) \in \mathbb{R}^2$. If F_{SP} and F_{VIX} are continuous, the copula C is unique and fully characterises the dependence structure, while the marginals are treated separately. In what follows we focus on estimating C for the pair $(r_t^{\text{SP}}, r_t^{\text{VIX}})$.

10.2 Tail Dependence Functionals

To quantify tail behaviour in a way that is comparable across different copula models, we focus on two *tail dependence indices* (TDIs) and a *joint exceedance probability* (JP). Throughout, let u and v denote the pseudo observations (probability integral transforms) for VIX and S&P 500 returns, respectively, so that each marginal is approximately $\text{Unif}(0, 1)$ and tail events correspond to u or v being close to 0 or 1. For a small probability level $\alpha \in (0, 0.5)$ define

$$\text{TDI}_1(\alpha) = \mathbb{P}(u < \alpha \mid v > 1 - \alpha), \quad \text{TDI}_2(\alpha) = \mathbb{P}(u > 1 - \alpha \mid v < \alpha),$$

and

$$\text{JP}(\alpha) = \mathbb{P}(u < \alpha, v > 1 - \alpha).$$

The indices $\text{TDI}_1(\alpha)$ and $\text{TDI}_2(\alpha)$ are *conditional* tail probabilities: they measure how likely one variable is to be in an extreme tail *given* that the other variable is simultaneously in its opposite tail. In our application, $\text{TDI}_1(\alpha)$ measures the probability that VIX is in its *lower* tail when the S&P 500 is in its *upper* tail, while $\text{TDI}_2(\alpha)$ measures the probability that VIX is in its *upper* tail when the S&P 500 is in its *lower* tail. Hence, larger values of $\text{TDI}_2(\alpha)$ indicate stronger “crash–spike” dependence, i.e., a tighter association between equity drawdowns and volatility surges in the joint tails.

The quantity $\text{JP}(\alpha)$ is an *unconditional* joint tail probability: it captures the overall frequency of a “crash–spike” configuration, namely VIX in its lower tail and the S&P 500 in its upper tail as defined (or, if u and v are defined in the opposite order in implementation, the corresponding crash–spike event consistent with that ordering). In particular, $\text{JP}(\alpha)$ provides a scale for how rare the joint tail event is, while the TDIs describe the strength of tail dependence *conditional* on being in the relevant tail of the conditioning variable.

Empirical values of these functionals are obtained by sample analogues based on the pseudo observations and correspond to functionals of the empirical copula C_T . For each parametric and nonparametric model we compute model implied values of the same statistics, either analytically from the fitted copula C or by Monte Carlo simulation when closed forms are not available.

10.3 Parametric Copula Benchmarks

We next consider several standard parametric copula families that are widely used in financial applications. These models provide interpretable summary parameters for dependence and facilitate model comparison via likelihood based criteria and tail dependence functionals.

10.3.1 Elliptical Copulas: Gaussian, Student- t , and Skew- t

The Gaussian copula is obtained by applying the multivariate normal distribution to Gaussianised versions of (u, v) . Let Φ^{-1} denote the inverse standard normal CDF and $\Phi_2(\cdot, \cdot; \rho)$ the bivariate normal CDF with correlation parameter $\rho \in (-1, 1)$. The Gaussian copula with parameter ρ is defined by

$$C_\rho^{\text{Gauss}}(u, v) = \Phi_2(\Phi^{-1}(u), \Phi^{-1}(v); \rho).$$

The corresponding copula density $c_\rho^{\text{Gauss}}(u, v)$ depends only on ρ and implies a symmetric dependence structure. The Gaussian copula does not exhibit tail dependence in the strict sense: the coefficient of tail dependence is equal to zero for any ρ with absolute value less than one. As a result, this model tends to underestimate the probability of joint extreme events when used with heavy tailed marginals.

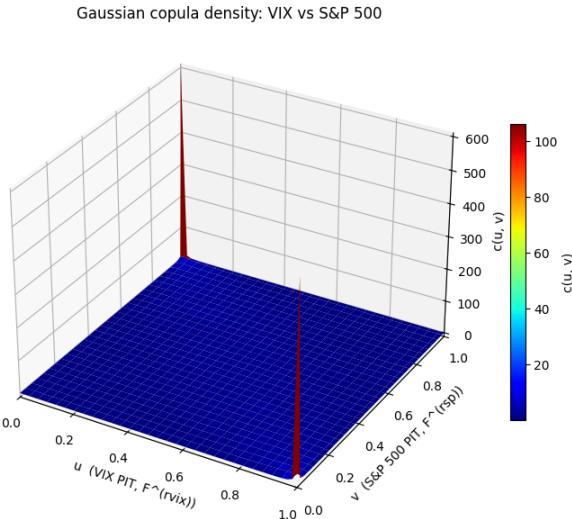


Figure 6: Gaussian copula density surface.

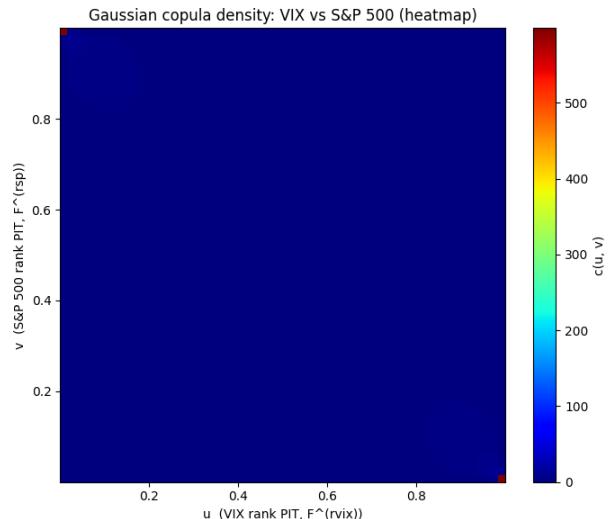


Figure 7: Gaussian copula density heatmap.

Gaussian copula density: VIX vs S&P 500 (99.5 pct clipped)

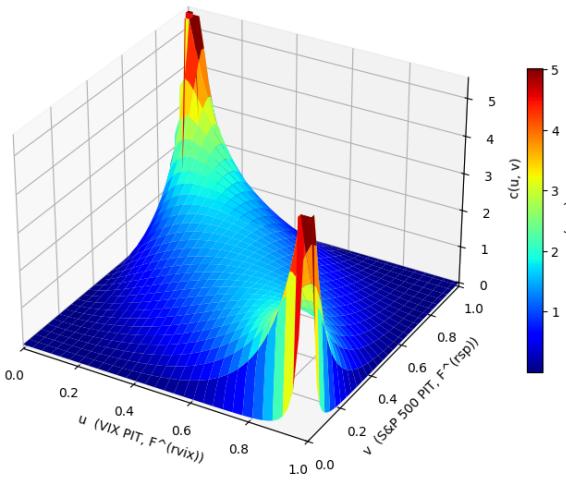


Figure 8: Gaussian copula density surface (clipped at the 99.5th percentile to improve visual interpretability).

Gaussian copula density: VIX vs S&P 500 (heatmap) 99.5 pct clipped

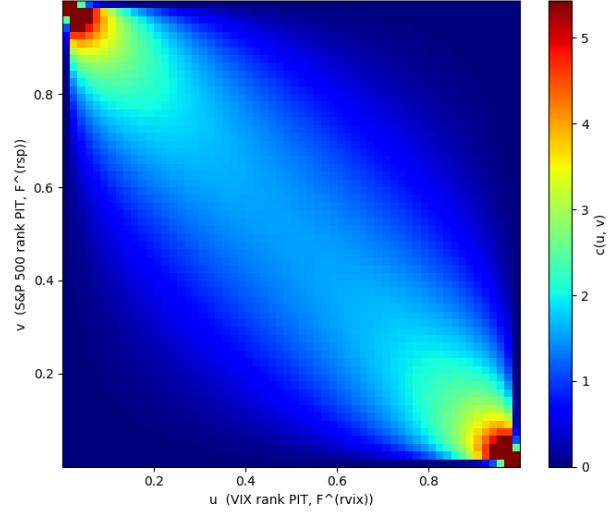


Figure 9: Gaussian copula density heatmap (clipped at the 99.5th percentile to improve visual interpretability).

Gaussian copula CDF

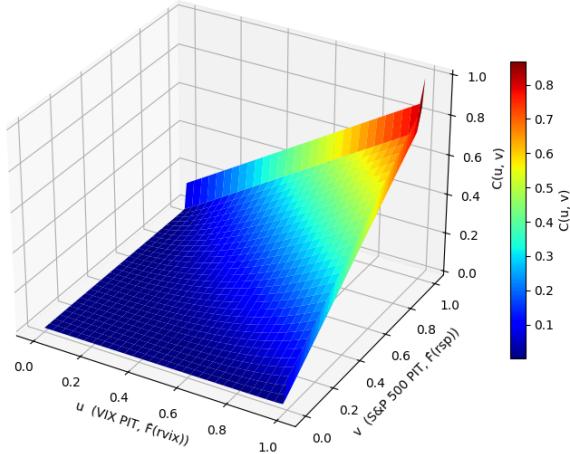


Figure 10: Gaussian copula cumulative density surface.

The Student- t copula generalises the Gaussian copula by replacing the normal kernel with a multivariate t distribution. Let T_ν^{-1} denote the inverse univariate Student- t CDF with $\nu > 2$ degrees of freedom and let $T_{2,\nu}(\cdot, \cdot; \rho)$ denote the bivariate t CDF. The t copula with correlation parameter ρ and degrees of freedom ν is given by

$$C_{\rho, \nu}^t(u, v) = T_{2,\nu}(T_\nu^{-1}(u), T_\nu^{-1}(v); \rho).$$

This copula retains symmetry but allows for nonzero tail dependence, with the strength of joint tail dependence increasing as ν decreases.

t copula density: VIX vs S&P 500

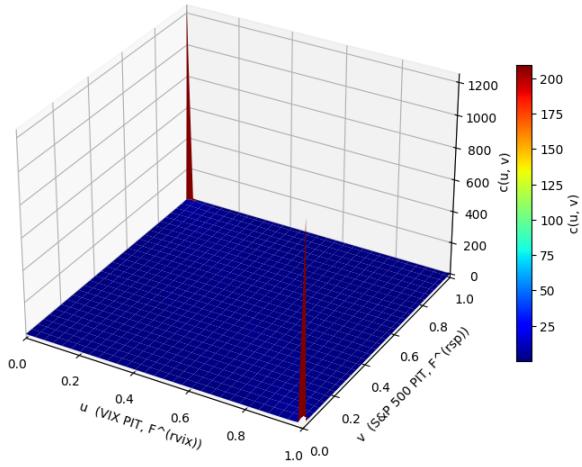


Figure 11: Student- t copula density surface.

t copula density: VIX vs S&P 500 (heatmap)

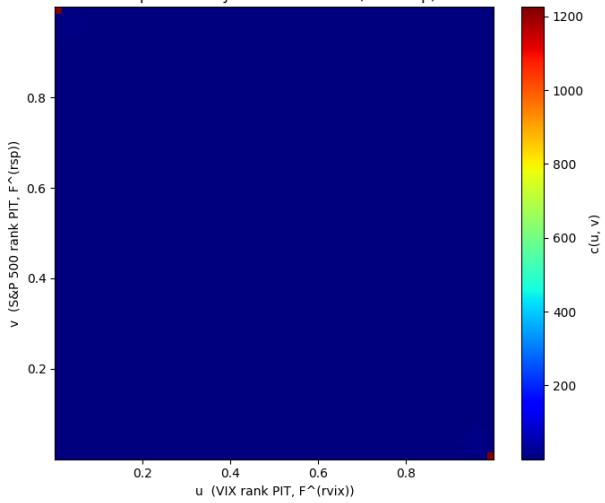


Figure 12: Student- t copula density heatmap.

t copula density: VIX vs S&P 500 (99.5 pct clipped)

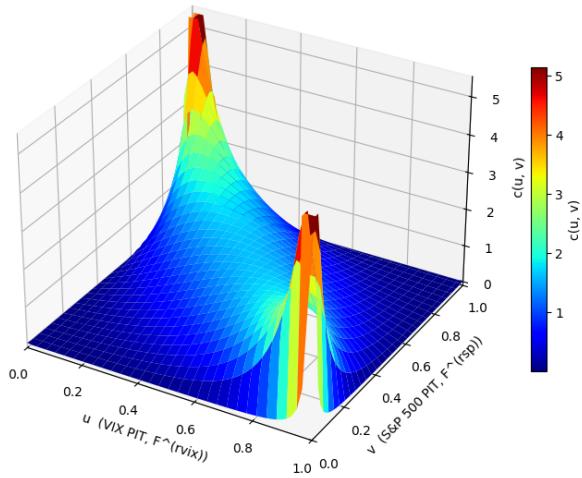


Figure 13: Student- t copula density surface (clipped at the 99.5th percentile to improve visual interpretability).

t copula density: VIX vs S&P 500 (heatmap) 99.5 pct clipped

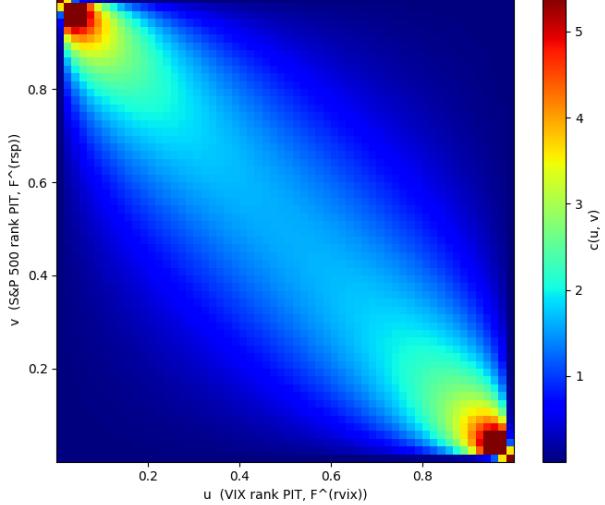


Figure 14: Student- t copula density heatmap (clipped at the 99.5th percentile to improve visual interpretability).

t copula CDF

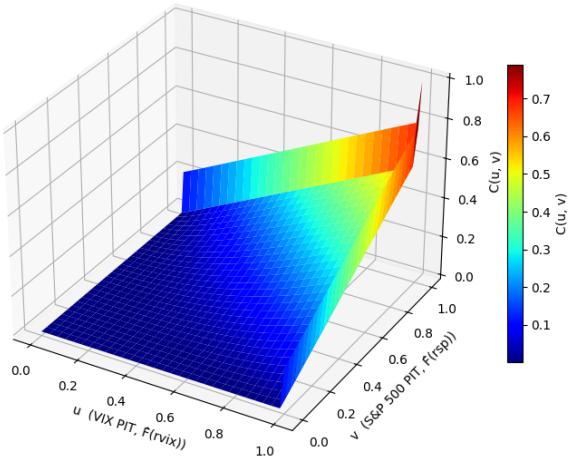


Figure 15: Student- t cumulative copula density heatmap.

To allow for potential asymmetry between the joint lower and upper tails we also consider a skew- t copula. Conceptually, this is obtained by replacing the symmetric t kernel with a bivariate skew- t distribution, so that

$$C_{\rho,\nu,\alpha}^{\text{Skew-}t}(u, v) = F_{2,\nu,\alpha}^{\text{Skew-}t}(Q_{\nu,\alpha}^{\text{Skew-}t}(u), Q_{\nu,\alpha}^{\text{Skew-}t}(v); \rho),$$

where $F_{2,\nu,\alpha}^{\text{Skew-}t}$ denotes a bivariate skew- t CDF with correlation ρ , degrees of freedom ν and skewness parameter α , and $Q_{\nu,\alpha}^{\text{Skew-}t}$ is the corresponding univariate skew- t quantile function. The additional parameter α tilts probability mass towards one tail, generating asymmetric tail dependence; setting $\alpha = 0$ recovers the symmetric t copula as a special case.

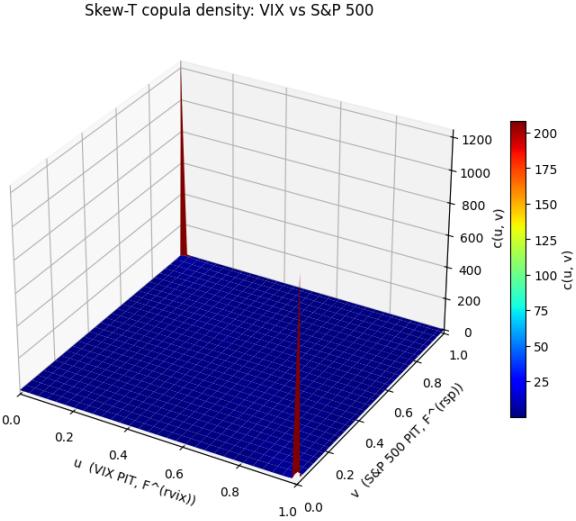


Figure 16: Skew- t copula density surface.

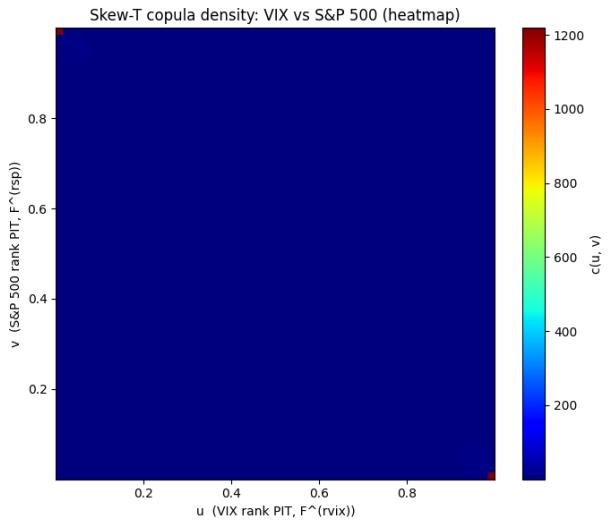


Figure 17: Skew- t copula density heatmap.

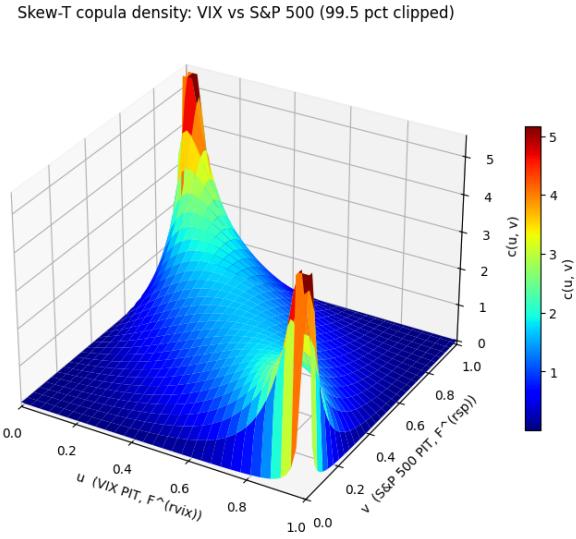


Figure 18: Skew- t copula density surface (clipped at the 99.5th percentile to improve visual interpretability).

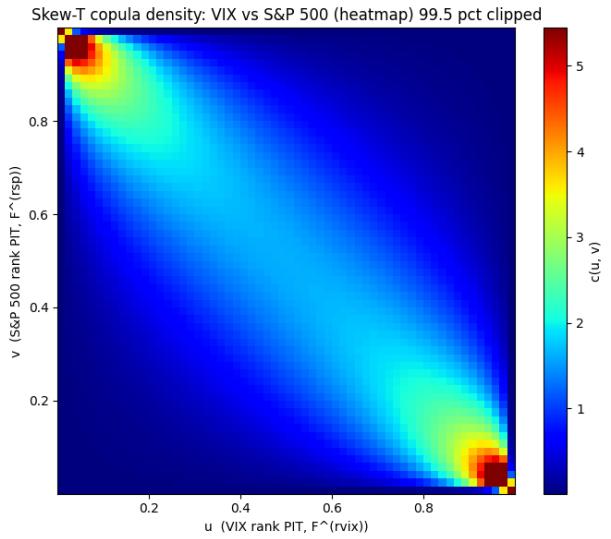


Figure 19: Skew- t copula density heatmap (clipped at the 99.5th percentile to improve visual interpretability).

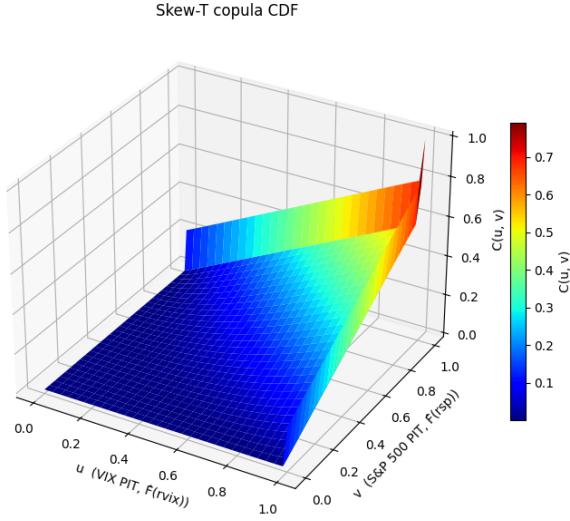


Figure 20: Skew- t copula cumulative density heatmap.

Parameters for all three elliptical copulas are estimated by maximising the copula log likelihood based on pseudo observations,

$$\hat{\theta} = \arg \max_{\theta} \sum_{t=1}^T \log c_{\theta}(u_t, v_t),$$

where c_{θ} denotes the copula density under the chosen family and θ collects (ρ) for the Gaussian copula, (ρ, ν) for the t copula, and (ρ, ν, α) for the skew- t copula. Model fit is assessed using the maximised log likelihood and information criteria such as AIC and BIC, and the implied tail indices $TDI_1(\alpha)$, $TDI_2(\alpha)$, and $JP(\alpha)$ are obtained by Monte Carlo simulation from the fitted copula.

10.3.2 Archimedean Copulas and Rotations

To allow for asymmetric tail dependence we also consider copulas from the Archimedean class. An Archimedean copula is defined by a generator function $\varphi : [0, 1] \rightarrow [0, \infty]$ that is continuous, strictly decreasing, and satisfies $\varphi(1) = 0$. The copula is then given by

$$C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)),$$

where φ^{-1} is the pseudo inverse of φ . Different choices of φ produce different dependence structures and tail patterns.

In our application we estimate several one parameter Archimedean families and their rotated versions:

- **Clayton** copula, with generator $\varphi(t) = t^{-\theta} - 1$ for $\theta > 0$, exhibits lower tail dependence. A 270° rotation creates strong dependence in the “crash spike” quadrant, where u is small and v is large. We also consider the 90° rotated Clayton, which places dependence in the opposite off diagonal quadrant; including both rotations provides a simple robustness check on whether the data favour crash spike dependence versus the reverse configuration.
- **Gumbel** copula, with generator $\varphi(t) = (-\log t)^{\theta}$ for $\theta \geq 1$, exhibits upper tail dependence. A 270° rotation again realigns the upper tail dependence to match the empirical configuration.

- **Joe** copula, defined by $\varphi(t) = -\log(1 - (1 - t)^\theta)$ for $\theta \geq 1$, captures strong upper tail dependence; a 270° rotation concentrates dependence in the “crash–spike” region.
- **Frank** copula, defined for $\theta \in \mathbb{R}$, is symmetric and does not exhibit tail dependence; as $\theta \rightarrow 0$ it converges to the independence copula.

Parameters are estimated by pseudo maximum likelihood as for the elliptical copulas. For the Archimedean families and their rotations, we exploit analytical expressions for the copula CDF to compute $\text{TDI}_1(\alpha)$, $\text{TDI}_2(\alpha)$, and $\text{JP}(\alpha)$ directly from the fitted copula.

Across the Archimedean specifications, the density surfaces and heatmaps reveal how strongly each copula concentrates probability mass into particular joint tail configurations of VIX and S&P 500 returns. For the rotated Clayton copulas, dependence is highly localised: the 270° rotation produces an intense spike concentrated in a single off diagonal corner of the unit square (Figures 23–24), while the 90° rotation flips this concentration to the opposite off diagonal corner (Figures 28–29). Economically, such “corner-dominated” densities imply that extreme co movements are driven by one dominant regime: one variable is very likely to be in an extreme tail precisely when the other sits in the opposite tail. In the S&P 500–VIX context, this corresponds to a highly asymmetric market state in which sharp equity drawdowns are tightly linked to large volatility spikes (or, under the opposite rotation, the reverse joint extreme), with relatively little probability mass allocated to other joint outcomes.

The rotated Gumbel and Joe copulas display a different geometry. Rather than a single corner dominating the entire fit, both show a broader ridge of elevated density running along the off diagonal direction with additional corner intensification (Figures 33–34 and 38–39). This shape suggests a more gradual form of negative dependence: as the S&P 500 shifts toward its lower tail, the VIX tends to shift toward its upper tail over a wider range of severities, not only at the most extreme points. In real terms, these copulas are consistent with the idea that volatility tends to rise as equities fall in a relatively continuous manner, while still allowing particularly high likelihood for joint extremes during major stress episodes.

Finally, the Frank copula yields the most spread out pattern among the Archimedean models considered here (Figures 43–44). Its density is spread broadly across the unit square, with less localised corner spikes, and it remains visually interpretable even without aggressive clipping, reflecting its lack of tail concentration relative to the other families. Substantively, this corresponds to a dependence structure that is closer to “average” association throughout the distribution: it can represent overall negative comovement between VIX and the S&P 500, but it does not attribute a disproportionate share of probability to specific joint tail regimes. Hence, compared with rotated Clayton, Gumbel, and Joe, the Frank copula implies a more spread out and less tail-driven linkage between equity selloffs and volatility spikes.

The cumulative copula surfaces provide a complementary view by showing how probability mass accumulates over (u, v) . When the density is highly concentrated in a corner, the corresponding CDF surface exhibits a steep “wall” in that region (Figures 25 and 30), whereas the Gumbel, Joe, and Frank specifications display more gradual transitions in their CDFs that mirror their comparatively smoother density patterns (Figures 35, 40, and 45).

Clayton copula (270 deg) density: VIX vs S&P 500

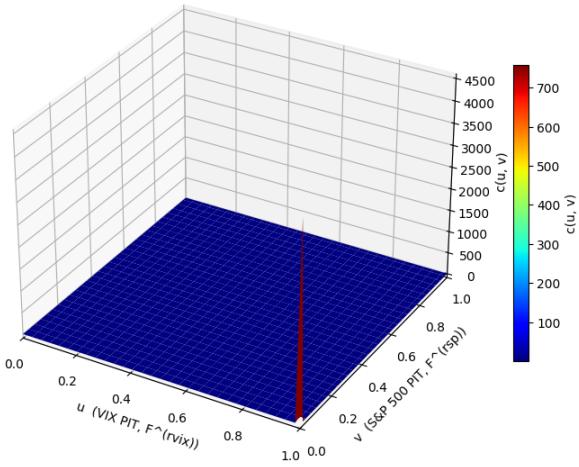


Figure 21: Clayton (270°) copula density surface.

Clayton copula (270 deg) density: VIX vs S&P 500 (heatmap)

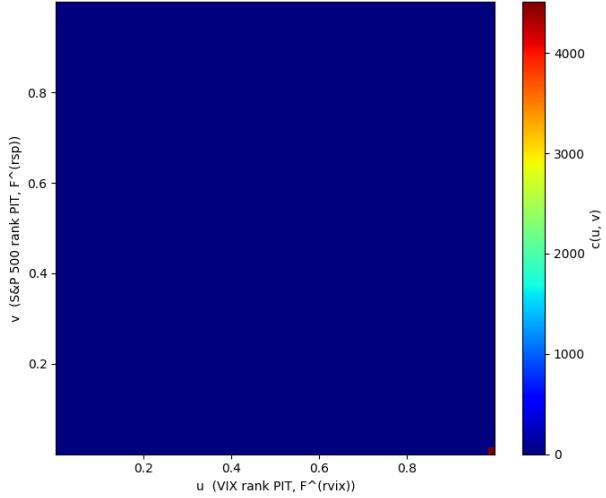


Figure 22: Clayton (270°) copula density heatmap.

Clayton copula (270 deg) density: VIX vs S&P 500 (99.5 pct clipped)

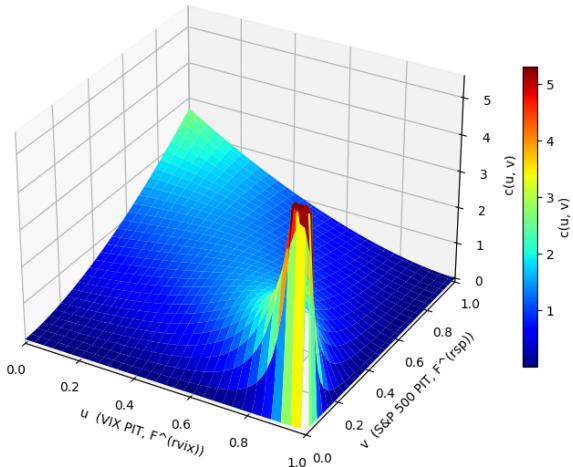


Figure 23: Clayton (270°) copula density surface (clipped at the 99.5th percentile to improve visual interpretability).

Clayton copula (270 deg) density: VIX vs S&P 500 (heatmap) 99.5 pct clipped

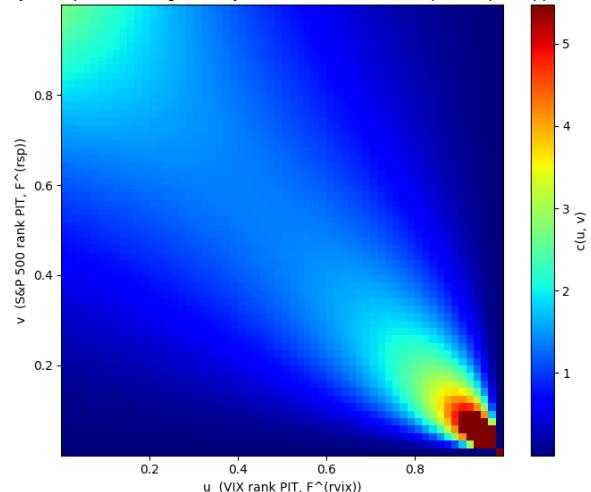


Figure 24: Clayton (270°) copula density heatmap (clipped at the 99.5th percentile to improve visual interpretability).

Clayton copula (270 deg) CDF

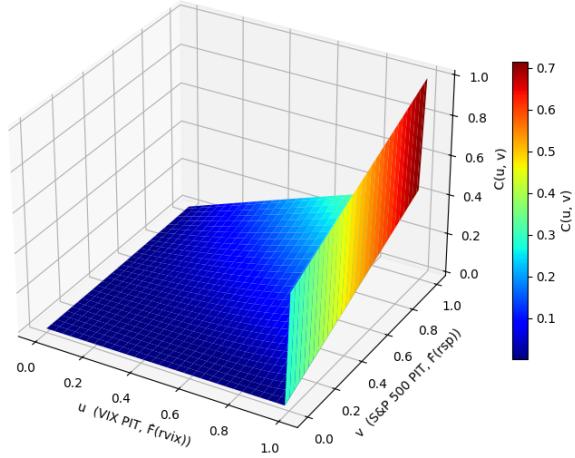


Figure 25: Clayton (270°) cumulative copula density heatmap.

Clayton copula (90 deg) density: VIX vs S&P 500

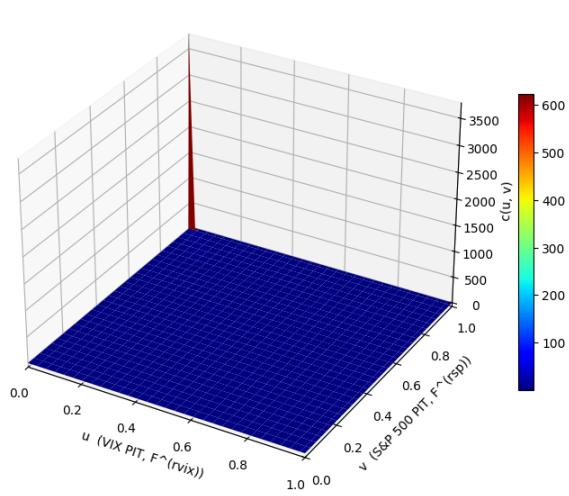


Figure 26: Clayton (90°) copula density surface.

Clayton copula (90 deg) density: VIX vs S&P 500 (heatmap)

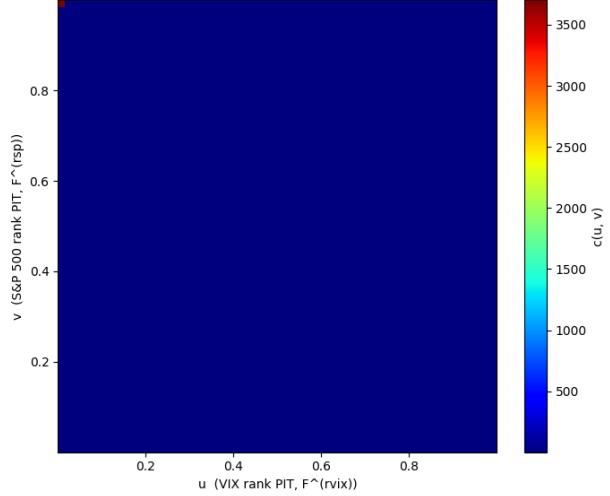


Figure 27: Clayton (90°) copula density heatmap.

Clayton copula (90°) density: VIX vs S&P 500 (99.5 pct clipped)

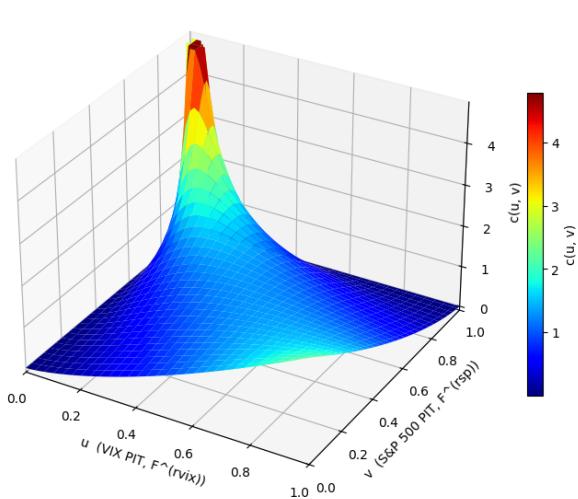


Figure 28: Clayton (90°) copula density surface (clipped at the 99.5th percentile to improve visual interpretability).

Clayton copula (90°) density: VIX vs S&P 500 (heatmap) 99.5 pct clipped

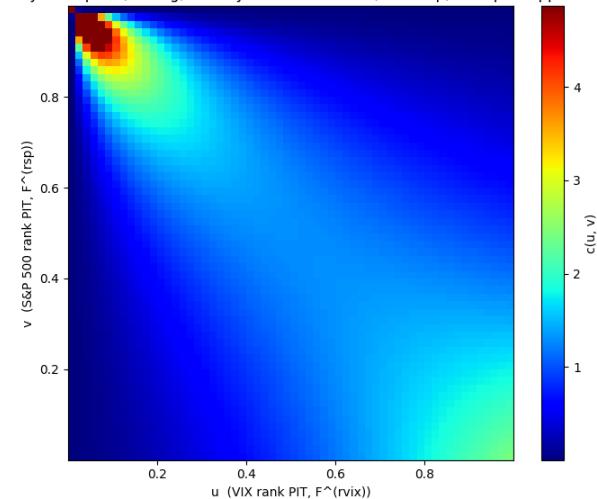


Figure 29: Clayton (90°) copula density heatmap (clipped at the 99.5th percentile to improve visual interpretability).

Clayton copula (90°) CDF

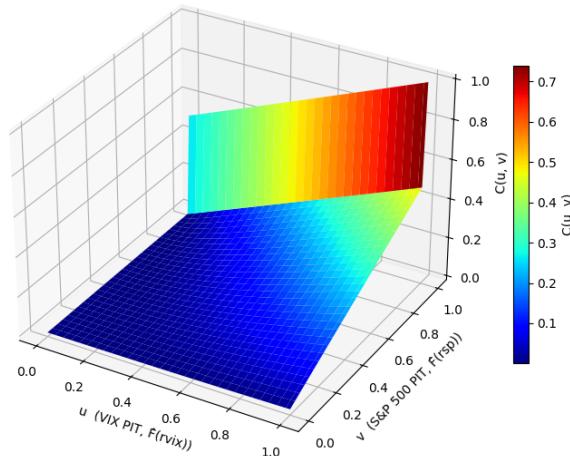


Figure 30: Clayton (90°) cumulative copula density heatmap.

Gumbel copula (270 deg) density: VIX vs S&P 500

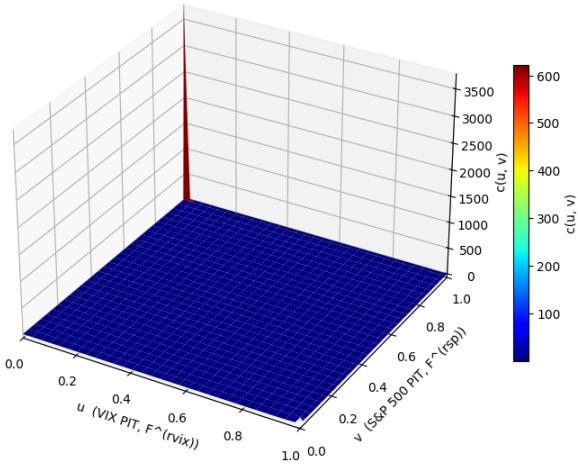


Figure 31: Gumbel (270°) copula density surface.

Gumbel copula (270 deg) density: VIX vs S&P 500 (heatmap)

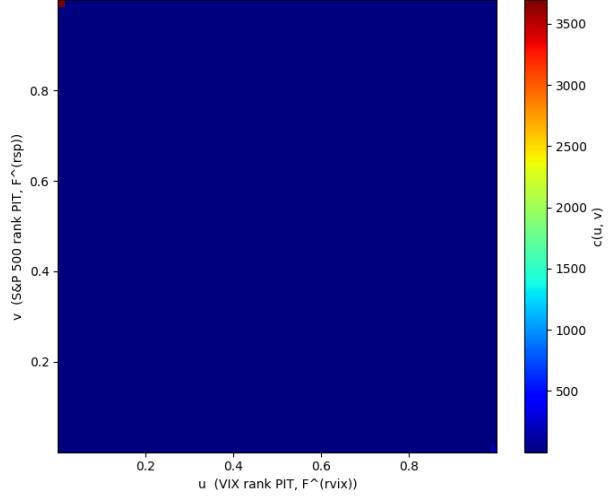


Figure 32: Gumbel (270°) copula density heatmap.

Gumbel copula (270 deg) density: VIX vs S&P 500 (99.5 pct clipped)

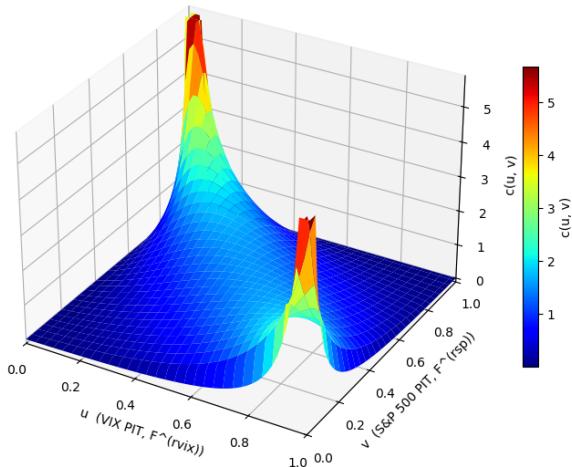


Figure 33: Gumbel (270°) copula density surface (clipped at the 99.5th percentile to improve visual interpretability).

Gumbel copula (270 deg) density: VIX vs S&P 500 (heatmap) 99.5 pct clipped

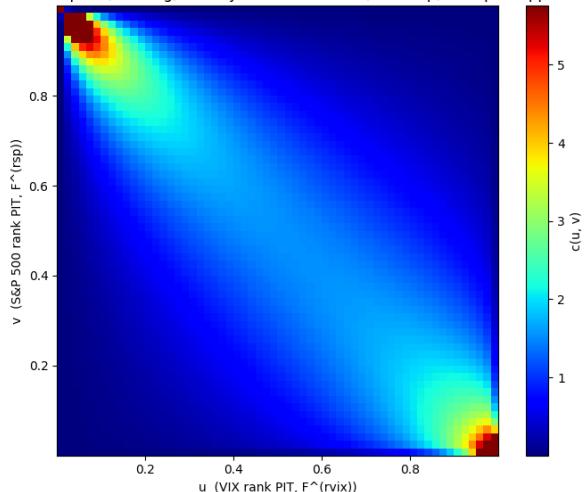


Figure 34: Gumbel (270°) copula density heatmap (clipped at the 99.5th percentile to improve visual interpretability).

Gumbel copula (270 deg) CDF

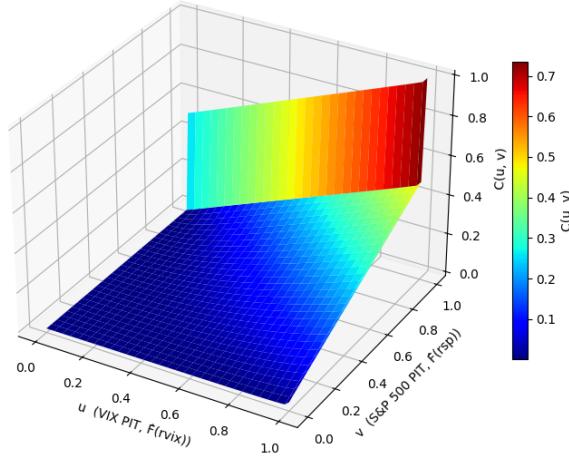


Figure 35: Gumbel (270°) cumulative copula density heatmap.

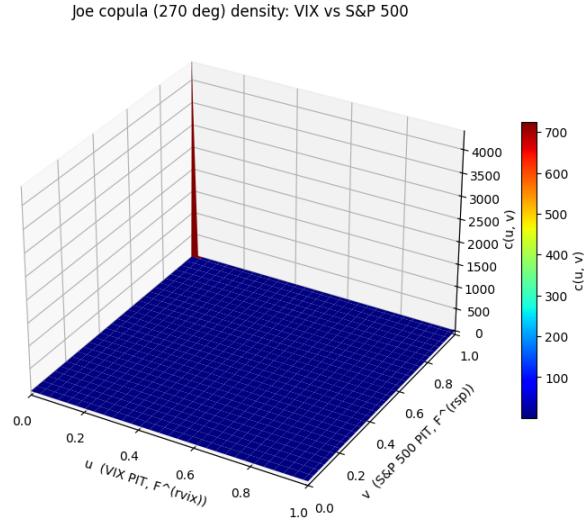


Figure 36: Joe (270°) copula density surface.

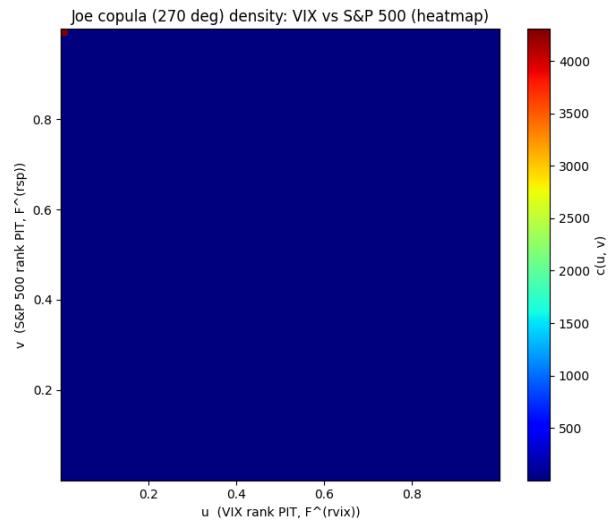


Figure 37: Joe (270°) copula density heatmap.

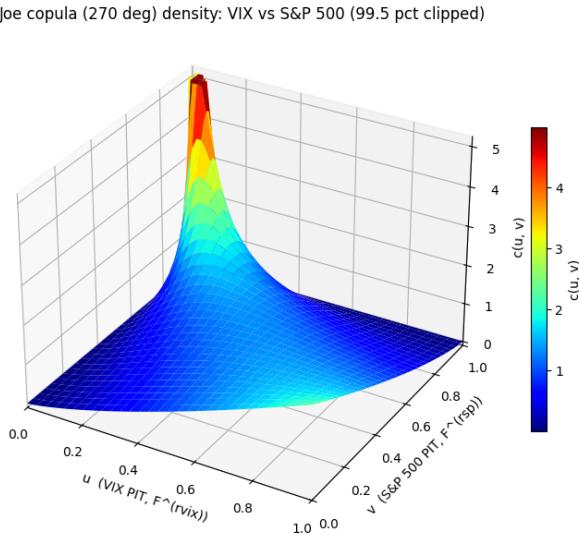


Figure 38: Joe (270°) copula density surface (clipped at the 99.5th percentile to improve visual interpretability).

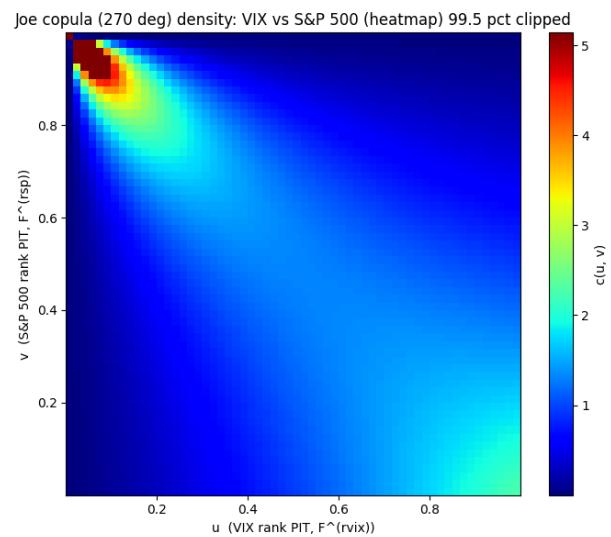


Figure 39: Joe (270°) copula density heatmap (clipped at the 99.5th percentile to improve visual interpretability).

Joe copula (270 deg) CDF

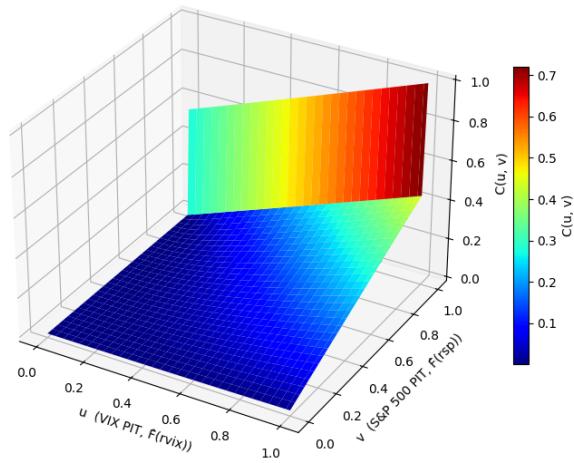


Figure 40: Joe (270°) cumulative copula density heatmap.

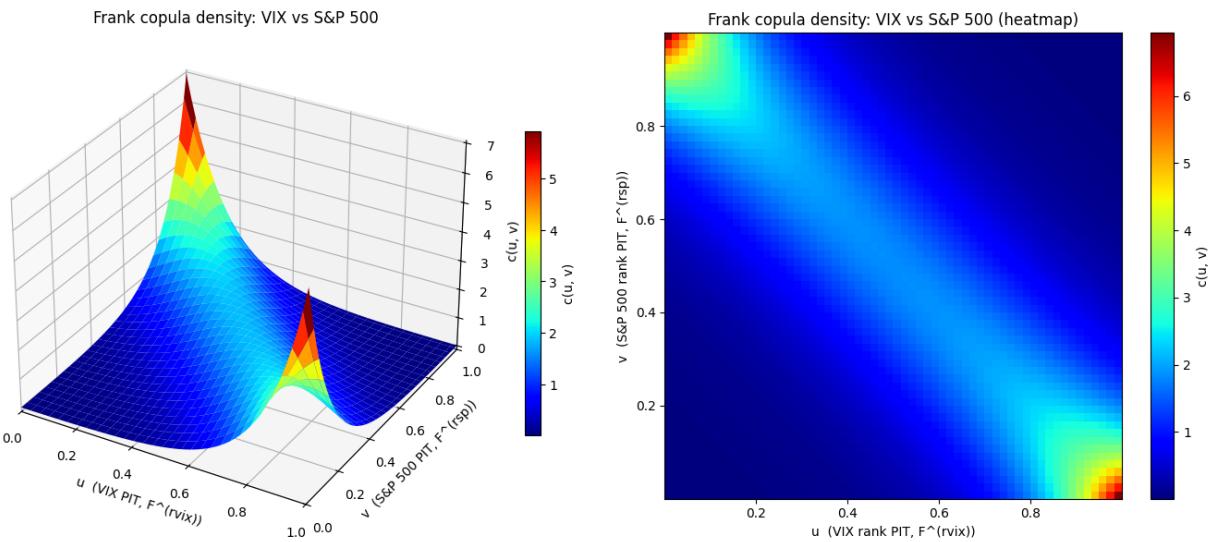


Figure 41: Frank copula density surface.

Figure 42: Frank copula density heatmap.

Frank copula density: VIX vs S&P 500 (99.5 pct clipped)

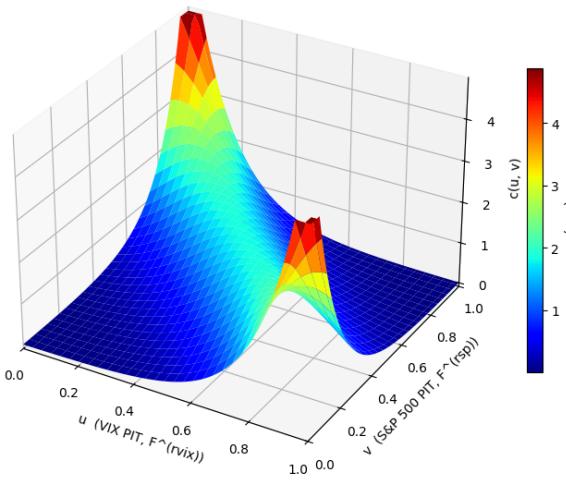


Figure 43: Frank copula density surface (clipped at the 99.5th percentile to improve visual interpretability).

Frank copula density: VIX vs S&P 500 (heatmap) 99.5 pct clipped

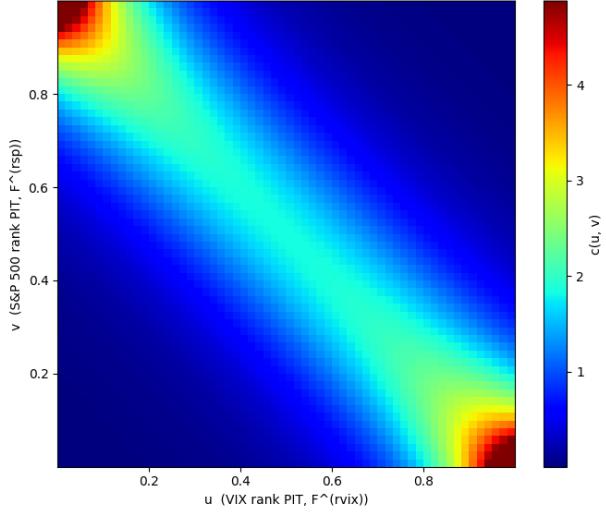


Figure 44: Frank copula density heatmap (clipped at the 99.5th percentile to improve visual interpretability).

Frank copula CDF

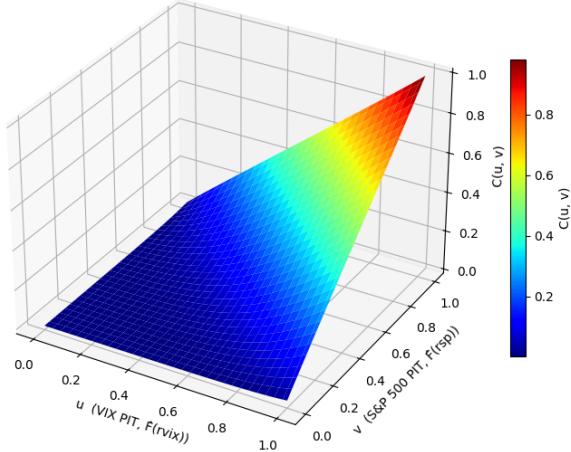


Figure 45: Frank cumulative copula density heatmap.

10.4 Exponential Series Estimator Copula

10.4.1 Model Specification

To obtain a flexible, fully nonparametric estimate of the copula density on $[0, 1]^2$, we adopt an Exponential Series Estimator (ESE) of (Wu, 2007). The ESE approximates the *log* density by a truncated series expansion and then exponentiates the approximation, which guarantees that the estimated density is everywhere nonnegative.

Let $(u_t, v_t) \in (0, 1)^2$ denote the pseudo-observations (rank based probability integral transforms). We model the copula density as an exponential family on the unit square:

$$c(u, v; \theta) = \exp\left(\sum_{(i,j) \in \mathcal{M}_m} \theta_{ij} \phi_{ij}(u, v) - \psi(\theta)\right), \quad (u, v) \in (0, 1)^2,$$

where $\phi_{ij}(u, v)$ are basis functions and $\theta = \{\theta_{ij}\}$ are coefficients. Following (Wu, 2007), we take orthonormal polynomial bases on $[0, 1]$ (e.g., normalized Legendre polynomials) for numerical stability, and form bivariate bases by products:

$$\phi_{ij}(u, v) = \phi_i(u)\phi_j(v).$$

The index set \mathcal{M}_m truncates the expansion to a finite dimension to control smoothness; consistent with standard ESE practice, we exclude the constant term so that the normalizing constant is identified:

$$\mathcal{M}_m = \{(i, j) : 0 < i + j < m\}, \quad m \geq 2.$$

The log partition (normalizing) function enforces the copula condition $\int_0^1 \int_0^1 c(u, v; \theta) du dv = 1$:

$$\psi(\theta) = \log \left(\int_0^1 \int_0^1 \exp \left(\sum_{(i,j) \in \mathcal{M}_m} \theta_{ij} \phi_{ij}(u, v) \right) du dv \right).$$

As m increases, the sieve space expands and the model can represent richer shapes on $[0, 1]^2$ (including asymmetric and locally concentrated dependence patterns), while maintaining a proper density via the exponential form (Wu, 2007).

10.4.2 Estimation Procedure

Given the pseudo observations $\{(u_t, v_t)\}_{t=1}^T$, the log likelihood of the ESE copula model is

$$\ell(\theta) = \sum_{t=1}^T \log c(u_t, v_t; \theta) = \sum_{t=1}^T \left(\sum_{0 < i+j < m} \theta_{ij} u_t^i v_t^j \right) - T \psi(\theta).$$

Direct evaluation of $\psi(\theta)$ is not available in closed form, so we approximate the integral over $[0, 1]^2$ numerically. Specifically, we construct a regular grid $\{(u_g, v_g)\}_{g=1}^G$ on the unit square and approximate

$$\psi(\theta) \approx \log \left(\frac{1}{G} \sum_{g=1}^G \exp \left(\sum_{0 < i+j < m} \theta_{ij} u_g^i v_g^j \right) \right).$$

The gradient of $\ell(\theta)$ with respect to each coefficient θ_{ij} can be derived analytically using the same grid approximation, which allows the use of efficient gradient based optimisation routines such as BFGS. In practice we initialise θ at zero, which corresponds to a uniform copula density, and iterate until convergence of the log likelihood. The degree parameter m and the grid resolution G are chosen to balance flexibility and numerical stability: small values of m risk underfitting the empirical copula, while excessively large values can lead to overfitting and erratic behaviour near the boundaries.

10.4.3 Estimated Copula Surface and Interpretation

Once the parameter vector $\hat{\theta}$ has been obtained, the estimated ESE copula density is

$$\hat{c}(u, v) = c(u, v; \hat{\theta}).$$

Visually, the ESE fit exhibits a pronounced high density ridge running along the off diagonal of the unit square, with the strongest mass concentrated in the two opposite corners. In particular, the density is largest in the bottom right corner (large u and small v), corresponding to the joint event that VIX is in its upper tail while the S&P 500 is in its lower tail. This is the characteristic “crash–spike” configuration. A second region of elevated density appears in the top left corner (small u and large v), consistent with the opposite “rally–crush” configuration in which equity returns are unusually strong while volatility is unusually low. Relative to these two regimes, the centre of the unit square carries substantially less density, indicating that extreme moves in one market are more likely to coincide with an offsetting move in the other market than with a similarly extreme move in the same direction.

Economically, these features align with the standard stress channel interpretation of the VIX: periods of sharp equity drawdowns coincide with volatility surges as market uncertainty and risk aversion rise, while strong equity rallies tend to occur alongside volatility compression as uncertainty resolves. Because the ESE model is not restricted to a one parameter Archimedean shape, it can capture this broad negative dependence pattern while allowing the concentration of probability mass to be uneven across the two off diagonal corners, reflecting the empirical asymmetry between downside stress episodes and upside relief episodes.

Individual coefficients $\hat{\theta}_{ij}$ are not easily interpretable in isolation, since they jointly determine the geometry of the surface. Accordingly, we interpret the fitted ESE copula primarily through functionals derived from $\hat{c}(u, v)$ —such as joint exceedance probabilities and conditional tail event probabilities—computed numerically on $[0, 1]^2$ by integration or Monte Carlo simulation. The results for these are in the next section.

Exponential Series Estimator Copula Density: VIX vs S&P 500

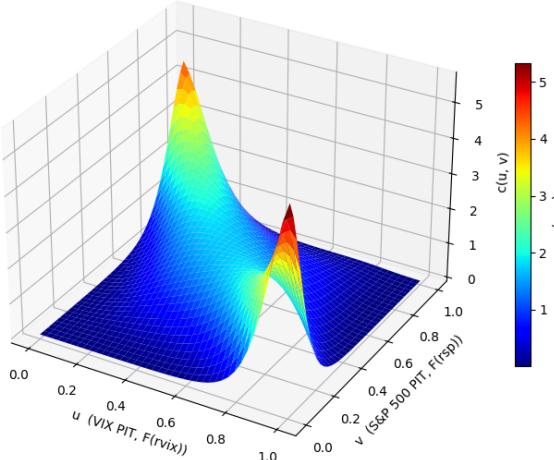


Figure 46: ESE copula density surface.

Exponential Series Estimator Copula Density: VIX vs S&P 500 (heatmap)

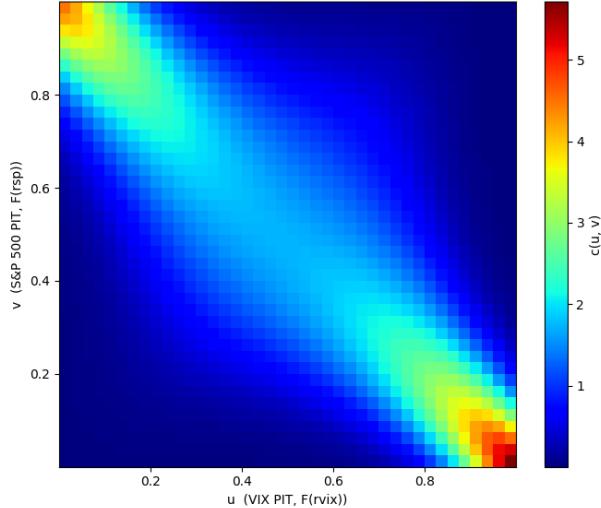


Figure 47: ESE copula density heatmap.

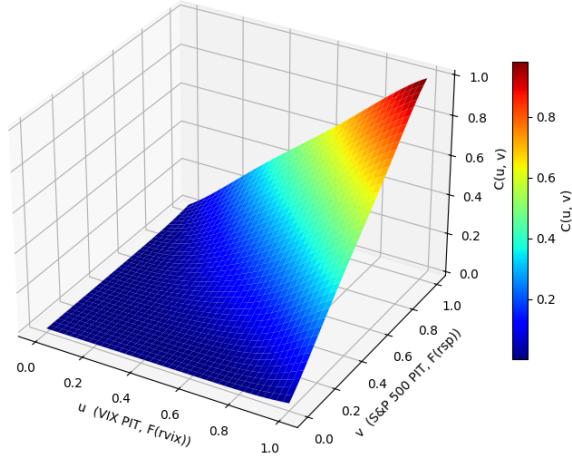


Figure 48: ESE cumulative copula density heatmap.

10.5 Comparison of Copula Models

To compare the ESE copula with the parametric elliptical and Archimedean specifications, it is useful to contrast both *shape restrictions* and the resulting fitted dependence patterns. The Gaussian and Student-*t* copulas impose an elliptical dependence structure: dependence is summarized by a correlation parameter (and in the *t* case a degrees of freedom parameter), which yields broadly symmetric dependence around the centre of the unit square and restricts how sharply probability mass can concentrate in specific off diagonal quadrants. In contrast, one parameter Archimedean families (and their rotations) can generate strong directional asymmetry, but typically do so by concentrating mass very strongly in a particular corner or along a narrow region, reflecting the limited flexibility of a single parameter generator.

The ESE copula sits between these extremes. Like the empirical (rank based) copula, it is non-parametric in the sense that it does not prespecify a particular copula family; however, unlike the empirical copula it produces a smooth density estimate and can be evaluated and integrated stably. Relative to the elliptical copulas, the ESE surface can reproduce localised features and asymmetric concentration that are visible in the data but are ruled out by elliptical symmetry. Relative to the rotated Archimedean copulas, the ESE fit can accommodate asymmetric tail behaviour without forcing all of the dependence into a single corner; instead it allows a broader ridge structure with varying intensity across the unit square, which is closer to the dependence patterns suggested by the heatmaps.

Quantitatively, we make this comparison using the same set of model evaluation objects across all specifications: log likelihood / information criteria for in sample fit, and common functionals of the copula such as joint exceedance probabilities and conditional tail event probabilities computed at $\alpha \in \{0.05, 0.01\}$. A copula that fits well should simultaneously (i) reproduce the observed off diagonal concentration in the density heatmap and (ii) deliver model implied tail event probabilities that are close to their empirical counterparts. In this sense, the ESE provides a flexible benchmark: it can be viewed as an approximation to the unknown “true” copula against which the parametric restrictions of elliptical and Archimedean families can be assessed. We investigate this further in the rest of this section.

Table 7: Summary statistics for copula models: log likelihood, information criteria, parameters, and tail dependence functionals.

	Empirical	Gaussian	Student- <i>t</i>	Skew- <i>t</i>	Clayton (270°)	Clayton (90°)	Frank	Gumbel (270°)	Joe (270°)	ESE
LogLik	—	3641	3716	3716	3116	2481	3496	3238	2362	3783
AIC	—	-7280	-7428	-7427	-6230	-4961	-6990	-6474	-4721	-7527
BIC	—	-7273	-7414	-7405	-6223	-4954	-6983	-6467	-4714	-7386
$\hat{\rho}$	—	-0.76	-0.76	-0.76	—	—	—	—	—	—
$\hat{\nu}$	—	—	10.00	10.39	—	—	—	—	—	—
$\hat{\theta}$	—	—	—	—	1.71	1.42	-6.94	2.05	2.27	—
TDI ₁ (0.05)	38.4	43.8	48.1	47.6	12.5	61.6	25.9	61.3	64.3	26.4
TDI ₁ (0.01)	22.4	32.2	40.3	41.3	2.7	61.3	6.5	60.1	64.3	15.0
TDI ₂ (0.05)	48.9	45.7	47.9	48.3	66.7	11.3	25.9	30.0	10.7	34.7
TDI ₂ (0.01)	35.3	34.0	39.2	41.7	66.6	2.4	6.5	15.8	2.2	18.8
JP(0.05)	1.9	2.2	2.4	2.4	0.6	3.1	1.3	3.1	3.2	1.0

The following paragraphs explain the findings and economic intuition of the results in Table 7.

Elliptical copulas: Gaussian, Student-*t*, and skew-*t*. Within the elliptical class, all three models estimate essentially the same correlation parameter, $\hat{\rho} \approx -0.76$, confirming a strong negative dependence between S&P 500 returns and VIX returns: equity drawdowns tend to coincide with volatility increases, while equity rallies tend to coincide with volatility compression. The key distinction across the elliptical copulas is therefore not the *sign* or *strength* of average dependence, but how much probability mass they allocate to joint extremes.

The Student-*t* copula achieves the best in sample fit among the three, with $\ell \approx 3716$ and the most favorable AIC and BIC (AIC ≈ -7428 , BIC ≈ -7414). The estimated degrees of freedom $\hat{\nu} = 10$ imply materially heavier joint tails than the Gaussian benchmark, meaning that extreme co movements (crash–spike episodes) occur more frequently than would be implied by an elliptical Gaussian dependence structure with the same $\hat{\rho}$. This is reflected in the tail functionals: relative to the Gaussian copula, the Student-*t* copula generates larger conditional tail probabilities and a higher joint crash probability, e.g. $JP(0.05) = 2.4$ versus 2.2 under Gaussian. Moreover, allowing for skewness via the skew-*t* copula has limited impact in our setting because copula estimation is performed on nonparametric rank based pseudo observations; the probability integral transform maps each margin to approximately $\text{Unif}(0, 1)$ and therefore removes marginal skewness by construction, leaving the skew-*t* parameter to capture only *copula level* asymmetry (which appears weak here). Compared to the empirical benchmarks, the Student-*t* copula is reasonably accurate for the economically relevant crash–spike direction: $\text{TDI}_2(0.05) = 47.9$ is close to the empirical 48.9, and at $\alpha = 0.01$ it remains close (39.2 versus empirical 35.3). However, it somewhat overstates the opposite conditional tail index $\text{TDI}_1(\alpha)$ (e.g. 48.1 vs 38.4 at $\alpha = 0.05$), indicating that the symmetric *t* copula tends to allocate too much mass to the mirror image extreme (high S&P 500 with low VIX) relative to what is observed in the data. Economically, this suggests that while heavy tails are important, the *asymmetry* of stress dynamics—crashes and volatility spikes being more salient than “rally–crush” episodes—is not fully captured by a symmetric elliptical specification.

The Gaussian copula, while sharing the same $\hat{\rho} \approx -0.76$, yields a substantially lower log likelihood ($\ell \approx 3641$) and worse information criteria. Its implied tail metrics are systematically smaller than those of the Student-*t* copula (e.g. $\text{TDI}_2(0.01) = 34.0$ versus 39.2 under Student-*t*), reflecting the fact that Gaussian dependence is not tail dependent and consequently underrepresents the frequency of joint extremes relative to heavy tailed alternatives. Notably, even the Gaussian copula slightly overstates the empirical joint crash probability at $\alpha = 0.05$ (2.2 vs 1.9), which is consistent with Gaussian dependence spreading mass smoothly over the off diagonal region rather than concentrating it in the empirically relevant corner; it captures negative association in the bulk but does not reproduce the observed sharp clustering during stress episodes.

The skew-*t* copula delivers essentially no improvement in fit over the Student-*t* copula (same $\ell \approx 3716$ and nearly identical AIC/BIC), with $\hat{\nu} \approx 10.39$ and tail functionals very close to the Student-*t* values. This indicates that, in the current implementation, the additional skewness parameter does not materially improve the copula level dependence fit. Two technical points are important for interpretation. First, the copulas are estimated on rank based pseudo observations, which remove marginal skewness and kurtosis by construction; any estimated “skewness” must therefore reflect genuine *dependence asymmetry* rather than marginal non normality. Second, if the

skew- t implementation introduces skewness primarily through ad hoc marginal like transformations (rather than a standard skew elliptical copula construction), the extra parameter may have limited ability to represent asymmetric dependence in (u, v) space and can effectively collapse back to the symmetric t fit. Taken together, the results in Table 7 suggest that a symmetric heavy tailed elliptical copula already captures most of the dependence that the data support, with residual asymmetry better handled by more flexible nonparametric approaches.

Archimedean copulas: rotated Clayton, rotated Gumbel, rotated Joe, and Frank. The one parameter Archimedean families impose strong shape restrictions that can generate pronounced directional asymmetry, especially after rotation. This is visible directly in the tail functionals.

For the rotated Clayton copulas, the tail behavior is extreme and one sided. Clayton (270°) produces very high crash–spike tail dependence in one direction, with $\text{TDI}_2(0.05) = 66.7$ and $\text{TDI}_2(0.01) = 66.6$, but extremely low values in the opposite direction ($\text{TDI}_1(0.05) = 12.5$, $\text{TDI}_1(0.01) = 2.7$). Its joint crash probability is far below the empirical benchmark, $\text{JP}(0.05) = 0.6$ versus 1.9, indicating that although the *conditional* probability can be very large once the conditioning tail event occurs, the model assigns too little *unconditional* mass to the relevant corner of the unit square. Clayton (90°) flips the pattern: TDI_1 becomes extremely large (61.6 and 61.3) while TDI_2 collapses (11.3 and 2.4), and $\text{JP}(0.05) = 3.1$ overshoots the empirical value. Both rotated Clayton variants therefore overconcentrate dependence into a single corner and fail to reproduce the more balanced empirical tail pattern, which is why their likelihoods are also substantially worse than the elliptical models (e.g. $\ell \approx 3116$ and $\ell \approx 2481$).

The rotated Gumbel and Joe copulas show a similar mismatch, but with dependence concentrated differently. Both imply very large $\text{TDI}_1(\alpha)$ (e.g. at $\alpha = 0.01$, $\text{TDI}_1 = 60.1$ for Gumbel and 64.3 for Joe) and comparatively small $\text{TDI}_2(\alpha)$ (e.g. 15.8 and 2.2 at $\alpha = 0.01$), while their $\text{JP}(0.05) = 3.1$ and 3.2 are materially above the empirical 1.9. This indicates that these rotated families can generate strong tail clustering, but in a way that places too much probability mass in a narrow tail region and not enough in the opposite tail configuration. Economically, these models exaggerate the extent to which tail events are “locked” into a single directional regime, whereas the data exhibit strong crash–spike behavior but still retain nontrivial mass across other joint outcomes.

Finally, the Frank copula provides a symmetric, tail independent Archimedean benchmark. Its estimated parameter $\hat{\theta} \approx -6.94$ reflects fairly strong overall negative dependence, and it achieves a competitive likelihood among the Archimedean models ($\ell \approx 3496$), though still below the elliptical leaders. However, its model implied tail functionals are far below the empirical values in both directions: $\text{TDI}_1(0.05) = \text{TDI}_2(0.05) = 25.9$ versus empirical 38.4 and 48.9, and $\text{JP}(0.05) = 1.3$ versus 1.9. This is consistent with tail independence: Frank can fit average negative association but systematically underestimates stress period co movements.

Nonparametric ESE copula. The ESE copula achieves the highest log likelihood ($\ell \approx 3783$) and the lowest AIC (≈ -7527), indicating that the additional flexibility is rewarded by a substantial increase in fit. Its BIC (≈ -7386) is worse than the Student- t copula’s BIC but remains relatively competitive, reflecting the trade off between flexibility and parsimony. In terms of tail functionals, ESE tracks the empirical crash–spike direction notably well at $\alpha = 0.05$: $\text{TDI}_2(0.05) = 34.7$ is below the empirical 48.9 but much closer than several Archimedean fits, while $\text{JP}(0.05) = 1.0$

remains conservative relative to 1.9. At $\alpha = 0.01$, $\text{TDI}_2(0.01) = 18.8$ is also conservative relative to 35.3. Overall, ESE appears to prioritize a smoother redistribution of mass across the unit square rather than producing the extreme corner concentration implied by the rotated one parameter Archimedean families.

Economically, the ESE results suggest that while crash–spike dependence is clearly present, a highly rigid one corner tail dependence structure is too extreme for the data; instead, the dependence is better described as strong negative association with elevated (but not degenerate) mass in stress configurations.

11 Limitations for Distribution and Dependence Analysis

We estimate copulas using nonparametric rank-based pseudo-observations (empirical PITs). This is robust and simple: it avoids committing to a parametric marginal model, is invariant to monotone transformations, and reduces sensitivity to outliers and heavy tails. The trade off is that rank based PITs remove marginal distributional features by construction, so characteristics such as marginal skewness, tail thickness, and time-varying scale are not modelled explicitly and can only appear indirectly through the copula.

A parametric alternative is to model the margins first (e.g. with Student- t /skew- t innovations and time varying volatility) and then apply parametric PITs to filtered residuals before fitting the copula. This can preserve and quantify marginal asymmetries (including skewness), but it is more complex for financial returns because of volatility clustering and conditional heteroskedasticity; credible parametric PITs typically require ARMA–GARCH-type filtering to obtain approximately i.i.d. innovations. This introduces additional modelling risk: misspecifying the volatility dynamics or innovation distribution can distort the residual PITs and bias the subsequent copula estimates. Hence we left this specification as a potential area for future work.

12 Conclusion for Distribution and Dependence Analysis

Three conclusions emerge from Table 7. First, dependence is strongly negative across all models, consistent with the role of the VIX as a market stress indicator and implying that large equity drawdowns tend to coincide with volatility spikes. Second, allowing for heavy tails through the Student- t copula materially improves fit and produces crash–spike tail metrics closer to the empirical benchmarks than the Gaussian copula, making it the most credible low dimensional parametric benchmark (especially under BIC). Third, rotated one-parameter Archimedean copulas can generate very strong directional tail concentration, but they typically misallocate mass by over emphasising one tail configuration and under emphasising the other, leading to tail probabilities that are either too extreme or too small. The ESE copula provides a flexible, smooth benchmark with the best in sample likelihood/AIC performance, and is useful for diagnosing which tail features are not well captured by rigid parametric families, though its tail event probabilities in Table 7 remain somewhat conservative relative to the empirical copula. Overall, these results from our particular sample and model setup motivate using the Student- t copula as a parsimonious dependence model and the ESE copula as a diagnostic reference point, while the regression analysis in the next section complements this distributional view by quantifying the conditional return response of the S&P 500 to VIX movements.

13 Regression Analysis

Related Literature In our regression models, we treat returns on VIX as a forward looking measure of market volatility following the literature in (Zhang et al., 2021). Relative to this literature, we do not study volatility forecasting directly. Instead, we take 2 key idea from (Zhang et al., 2021) which is treating the VIX as a model free implied volatility benchmark that contains forward looking information, and recognising the measurement error and simultaneity issues in regressions involving implied volatility. This motivates both our focus on the VIX as the main explanatory variable for S&P 500 returns and our use of instrumental variable methods (including lagged implied volatility) to address potential endogeneity.

Secondly, from (Sun and Wu, 2009), they mentioned asymmetry which suggests that the dependence is stronger in extreme bearish markets than in extreme bullish markets, and they tested this using tail dependence indices and a conditional dependence index that varies over the distribution of market returns. Then, they build simple regression models of VIX returns on S&P 500 returns, augmented with dummies for extreme tails, to capture this asymmetric tail behaviour. Our regression design is directly inspired by these findings: We begin with a simple linear model relating S&P 500 returns to VIX returns, then introduce interaction terms to test for state dependence (e.g. interactions with the sign of VIX returns, and with the sign of lagged S&P 500 returns).

Thirdly, our attempt to use the US Economic Policy Uncertainty (EPU) index as an instrument for VIX returns builds on the news based policy uncertainty literature. The EPU index (Baker et al., 2016) is constructed from US newspaper articles that mention economic terms, uncertainty, and policy related keywords; it is normalised by total article counts and updated using a rolling 30 day window. Conceptually, spikes in economic policy uncertainty could raise demand for protection, increase implied volatility and hence move VIX, suggesting that changes in EPU might be a relevant instrument for VIX returns. In the IV setup, we treat daily percentage changes in EPU as a potential source of exogenous variation in VIX, and test whether this satisfies the relevance and exogeneity conditions.

Methodology In this section, we study the relationship between S&P 500 returns on VIX returns using various regression models. Before estimating any models, the data were sorted in ascending order by date to preserve the time series structure and avoid any form of look ahead bias or data leakage. Next, we tested for stationarity for both `rsp` and `rvix` using the Augmented Dickey Fuller (ADF) test. The ADF tests strongly rejects the null hypothesis H_0 of a unit root, indicating that the return series are stationary. This justifies the application of standard linear time series methods on `rsp` and `rvix` without any additional differencing. Additionally, we conducted the Breusch Godfrey test using $P^{\frac{1}{3}}$ number of lags (where P is number of observations), to detect autocorrelation in the residuals of the regression model. The Breusch Godfrey test strongly rejects the null hypothesis H_0 of no autocorrelation. Therefore, to ensure our interpretation of statistical significance is reliable even in the presence of heteroskedasticity and residual autocorrelation, we report Newey West heteroskedasticity and autocorrelation consistent (HAC) standard errors and base all interpretation on the corresponding HAC robust t-statistics and p-values.

13.1 Linear & Non-Linear models

13.1.1 S&P 500 returns on VIX returns

We regress the returns of S&P 500 on VIX returns in interactions.R and the results are shown in Figure 49

$$r_t^{\text{SP}} = \beta_0 + \beta_1 r_t^{\text{VIX}} + e_t$$

```
> summary(linear_model)

Call:
lm(formula = rsp ~ rvix, data = df)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.4309 -0.3716 -0.0138  0.3627  8.9515 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.032645  0.008624   3.785 0.000155 ***
rvix        -0.122122  0.001268 -96.326 < 2e-16 ***  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7958 on 8513 degrees of freedom
Multiple R-squared:  0.5215,    Adjusted R-squared:  0.5215 
F-statistic: 9279 on 1 and 8513 DF,  p-value: < 2.2e-16

t test of coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  0.0326445  0.0077272   4.2246 2.418e-05 ***
rvix        -0.1221216  0.0050947 -23.9703 < 2.2e-16 ***  
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(a) Results of regressing S&P 500 on VIX returns

(b) HAC robust standard errors

Figure 49: OLS regression of S&P 500 returns on VIX returns with conventional and HAC-robust standard errors.

Interpretation We can see that the coefficient on VIX returns is -0.12, which means that for every 1 unit increase in VIX returns, S&P 500 returns decrease by -0.12 units. This linear model also has an R^2 of 0.52, meaning that the returns on VIX explains 52 % of the variations of S&P 500 returns. After accounting for serial correlation by computing HAC Robust SE shown in Figure 49b using $P^{\frac{1}{3}}$ number of lags (where P is number of observations), the coefficient is statistically significant.

A negative relationship between VIX and S&P 500 prices reflects that the VIX is a “fear gauge” of the equity market. When the VIX rises, it signals higher perceived uncertainty and risk, so investors avoid equities, causing prices to drop. Conversely, when the VIX falls and uncertainty subsides, equity prices tend to rise as demand for equities rise again.

13.1.2 S&P 500 returns on squared of VIX returns

To study whether the relationship between S&P 500 returns and VIX is non linear, we estimate a quadratic regression of daily S&P 500 returns, r_t^{SP} , on squared of VIX returns $r_t^{\text{VIX}} t$ in the form:

$$r_t^{\text{SP}} = \alpha + \beta_1 r_t^{\text{VIX}} + \beta_2 (r_t^{\text{VIX}})^2 + u_t \quad (1)$$

```

Call:
lm(formula = rsp ~ rvix + I(rvix^2), data = df)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.8100 -0.3679 -0.0057  0.3718  8.9694 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 1.509e-02 9.137e-03 1.652   0.0986 .  
rvix        -1.246e-01 1.337e-03 -93.215 < 2e-16 *** 
I(rvix^2)   3.792e-04 6.616e-05  5.731 1.03e-08 *** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 0.7943 on 8512 degrees of freedom
Multiple R-squared:  0.5234, Adjusted R-squared:  0.5232 
F-statistic: 4673 on 2 and 8512 DF, p-value: < 2.2e-16

t test of coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.01509321 0.01054895 1.4308 0.15253  
rvix        -0.12458724 0.00486278 -25.6206 < 2e-16 *** 
I(rvix^2)   0.00037917 0.00016174  2.3443 0.01908 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

```

(a) Results of regressing S&P 500 returns on squared of VIX returns

(b) HAC-robust standard errors

Figure 50: Quadratic regression

Interpretation Figure 50 shows that the linear VIX coefficient $\beta_1 \approx -0.125$ remains negative and statistically significant. This is similar to the magnitude of the linear model 49a, reaffirming the inverse relationship between VIX returns and S&P 500 returns. The quadratic regression attains an R^2 of 0.5232, implying that the model explains approximately 52.3% of the variation in r_t^{SP} in sample. Relative to the linear model, this represents only a modest improvement in fit.

The quadratic coefficient $\beta_2 \approx 0.00038$ is also positive and statistically significant, implying that the relationship is not purely linear. To characterize the non linear effect of VIX returns on S&P 500 returns, we can compute the marginal effect of VIX returns on S&P 500 returns as:

$$\frac{\partial r_t^{\text{SP}}}{\partial r_t^{\text{VIX}}} = -0.1246 + 2 \times 0.000379 r_t^{\text{VIX}}. \quad (2)$$

This marginal effect is negative over the range of r_t^{VIX} , implying that higher VIX returns are associated with lower S&P 500 returns. Moreover, because $\beta_2 > 0$, the marginal effect increases with r_t^{VIX} , meaning that the negative association becomes weaker when VIX returns are already high. Economically, this suggests a convex VIX returns relationship where volatility spikes coincide with equity selloffs, but the incremental effect of further increases in VIX returns attenuates during periods of very high market turbulence.

13.2 Interaction

In the same interactions.R file, we explore the use of 2 interaction terms. Firstly, we investigate if the sign of VIX returns has any effect on S&P 500 returns. In short, we find out if a negative VIX return will have a different effect in terms of magnitude compared to a positive VIX return on S&P 500 returns. We multiply VIX returns and an indicator variable (1 if VIX returns > 0 and 0 otherwise).

$$r_t^{\text{SP}} = \beta_0 + \beta_1 r_t^{\text{VIX}} + \beta_2 r_t^{\text{VIX}} \cdot \mathbb{1}_{\{r_t^{\text{VIX}} > 0\}} + e_t$$

$$\text{If } r_t^{\text{VIX}} \leq 0 : \quad \hat{r}_t^{\text{SP}} = \beta_0 + \beta_1 r_t^{\text{VIX}}$$

$$\text{If } r_t^{\text{VIX}} > 0 : \quad \hat{r}_t^{\text{SP}} = \beta_0 + (\beta_1 + \beta_2) r_t^{\text{VIX}}$$

```

Call:
lm(formula = rsp ~ rvix + rvix_dummy, data = df_dummy)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.4664 -0.3703 -0.0114  0.3622  8.9582 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.025847  0.012401  2.084   0.0372 *  
rvix       -0.123755  0.002488 -49.732  <2e-16 *** 
rvix_dummy  0.002817  0.003693   0.763   0.4456  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.7958 on 8512 degrees of freedom
Multiple R-squared:  0.5216,    Adjusted R-squared:  0.5214 
F-statistic:  4639 on 2 and 8512 DF,  p-value: < 2.2e-16

t test of coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.0258470  0.0156362  1.6530  0.09836 .  
rvix       -0.1237549  0.0055232 -22.4064 < 2e-16 *** 
rvix_dummy  0.0028172  0.0060638   0.4646  0.64224 
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

```

(a) Summary of Model

(b) HAC robust standard errors

Figure 51: OLS regression of S&P 500 returns on VIX returns and Interaction Term.

From Figure 51a we can see that the coefficient on the interaction term is 0.0028. This means that if VIX returns are positive, the negative effect VIX returns have on S&P 500 returns is $\beta_1 + \beta_2$. However, if VIX returns are negative, the effect of VIX returns on S&P 500 returns is just β_1 .

However, in Figure 52b, where the HAC robust standard errors are used, the interaction term is insignificant, which means that the sign of VIX returns likely has no effect on how VIX returns affect S&P 500 returns. This suggests that the magnitude of movement of VIX is more important than the sign of VIX returns in explaining S&P 500 returns.

Interaction between VIX returns and sign of lag values of S&P 500 returns Next, we test out another interaction term which is VIX returns multiplied by an indicator variable that is 1 if yesterday's returns on S&P 500 is positive and 0 otherwise. In economic sense, we investigate if the effect of return of VIX on S&P 500 returns is affected by yesterday's sign of return on S&P 500.

$$r_t^{\text{SP}} = \beta_0 + \beta_1 r_t^{\text{VIX}} + \beta_2 (r_t^{\text{VIX}} \cdot \mathbb{1}\{r_{t-1}^{\text{SP}} > 0\}) + e_t$$

```

Call:
lm(formula = rsp ~ rvix + rsp_dummy, data = df_interaction)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.4392 -0.3711 -0.0156  0.3606  8.9656 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.01855  0.01270   1.460   0.144    
rvix       -0.12223  0.00127 -96.263  <2e-16 *** 
rsp_dummy   0.02612  0.01733   1.507   0.132    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.7958 on 8511 degrees of freedom
Multiple R-squared:  0.5216,    Adjusted R-squared:  0.5215 
F-statistic:  4641 on 2 and 8511 DF,  p-value: < 2.2e-16

t test of coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.0185498  0.0114503  1.620   0.1053    
rvix       -0.1222268  0.0050887 -24.019 < 2e-16 *** 
rsp_dummy   0.0261165  0.0167090   1.563   0.1181    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

```

(a) Summary of Model

(b) HAC robust standard errors

Figure 52: OLS regression of S&P 500 returns on VIX returns and Interaction Term.

From Figure 52a we can see that the coefficient of the interaction term is 0.026. This means that if yesterday's returns on S&P 500 was positive, returns on VIX has a weaker negative effect on returns on S&P 500. Intuitively this means that after the market has done well, investors are slightly less reactive to VIX moves.

However, after computing the HAC robust standard errors in Figure 52, we can see that the coefficient on the interaction term is statistically insignificant even at a 10% significance level.

Interaction between VIX returns and quantiles of VIX returns Next, similar to (Sun and Wu, 2009), where they investigated if quantiles of returns on S&P 500 had any effect on VIX returns. In our case, we investigate if the quantiles of VIX returns has any effect on returns on S&P 500.

$$r_t^{\text{SP}} = \beta_0 + \beta_1 r_t^{\text{VIX}} + \beta_2 (r_t^{\text{VIX}} \cdot 1\{r_t^{\text{VIX}} > Q_{95}\}) + \beta_3 (r_t^{\text{VIX}} \cdot 1\{r_t^{\text{VIX}} < Q_5\}) + \varepsilon_t$$

Where:

- $1\{r_t^{\text{VIX}} > Q_{95}\}$ is an indicator function that equals 1 if r_t^{VIX} is above the 95th percentile Q_{95} of VIX returns, and 0 otherwise,
- $1\{r_t^{\text{VIX}} < Q_5\}$ is an indicator function that equals 1 if r_t^{VIX} is below the 5th percentile Q_5 of VIX returns, and 0 otherwise,

```
Call:
lm(formula = rsp ~ rvix + is_upper + is_lower, data = df_quantile)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.6402 -0.3708 -0.0145  0.3705  8.9577 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.027334  0.009027  3.028  0.00247 **  
rvix        -0.128985  0.002022 -63.782 < 2e-16 ***  
is_upper     0.012865  0.002918  4.409 1.05e-05 ***  
is_lower     0.008581  0.003407  2.519  0.01179 *    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 0.795 on 8511 degrees of freedom
Multiple R-squared:  0.5227, Adjusted R-squared:  0.5225 
F-statistic: 3106 on 3 and 8511 DF,  p-value: < 2.2e-16

t test of coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 0.0273342  0.0088008  3.1059  0.001903 **  
rvix        -0.1289851  0.0038526 -33.4799 < 2.2e-16 ***  
is_upper     0.0128653  0.0069674  1.8465  0.064856 .  
is_lower     0.0085806  0.0067522  1.2708  0.203835 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(a) Summary of Model

(b) HAC robust standard errors

Figure 53: OLS regression of S&P 500 returns on VIX returns and Interaction Term

From Figure 53a we can see that the coefficient of both β_2 & β_3 are positive, which means that if VIX returns belong in the 95th quantile or 5th quantile, the negative effect of VIX returns on S&P 500 returns would decrease. However, if we look at the HAC robust standard errors in 53b, we see that β_3 is insignificant. β_2 is significant on a 5% significance level. In economic sense, this means that an extremely high returns on VIX which indicates extreme fear, causing investors to step in and 'buying the dip' on S&P 500, pushing prices up slightly.

13.3 Autoregressive Model

In this section, we model daily S&P 500 returns, r_t^{SP} , using a univariate autoregressive (AR) model. A general AR(p) model for r_t^{SP} takes the form:

$$r_t^{\text{SP}} = \alpha_0 + \sum_{i=1}^p \alpha_i r_{t-i}^{\text{SP}} + \varepsilon_t \quad (3)$$

where p denotes the lag order. To choose optimal p , a grid search over $p \in 0, \dots, 10$ was conducted using Bayesian Information Criteria (BIC). BIC selects $p^* = 1$ as the optimal lag order. Hence, an AR(1) model for r_t^{SP} is adopted:

$$r_t^{\text{SP}} = 0.0329 - 0.0863 r_{t-1}^{\text{SP}} + \varepsilon_t. \quad (4)$$

```

z test of coefficients:
Estimate Std. Error z value Pr(>|z|)
ar1 -0.086291  0.010798 -7.9917 1.331e-15 ***
intercept 0.032906  0.011433  2.8780  0.004002 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
[1] 0.007330052

t test of coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.035679  0.012428  2.8710  0.004102 **
L(rsp, 1) -0.086299  0.010799 -7.9913 1.509e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

(a) Summary of model

(b) HAC robust standard Errors

Figure 54: AR(1) model for r_t^{SP}

Interpretation The intercept of 0.035679 means that, ignoring day to day fluctuations, AR(1) implies a small positive average daily return on the S&P 500 (Figure 54a). The r_{t-1}^{SP} coefficient is -0.0863 , which is negative and statistically significant (Figure 54). Economically, this indicates mean reversion. When the market has had a positive return the day before, it tends to pull back in the current period and vice versa. This suggests that the stock market might experience short term reversals, as large positive or negative returns are likely followed by opposing movements.

Although the magnitude of the coefficient (-0.0863) might seem small, given that the S&P 500 typically experiences modest daily fluctuations, a 1% change in yesterday's return can still lead to a significant shift in today's expected return, especially in a market with relatively low volatility. This result indicates that, while short term mean reversion exists, the market still exhibits a level of unpredictability, as returns are not fully predictable based on past performance alone.

The adjusted R squared of 0.00733 indicates that only 0.73% of the variation in daily S&P 500 is accounted for by the previous day's return. This suggests that the AR(1) model has limited explanatory power, implying that the stock market's daily movements are most likely influenced by other factors such as r_t^{VIX} .

13.4 Autoregressive Distributed Lag (ADL) Model

In this section, we explore the use of an autoregressive distributed lag (ADL) model to examine the relationship between r_t^{SP} and its lagged terms, and r_t^{VIX} and its lagged term. The general ADL(p, q) model is:

$$Y_t = \beta_0 + \sum_{i=1}^p \beta_i Y_{t-i} + \sum_{j=0}^q \delta_j X_{t-j} + \varepsilon_t \quad (5)$$

where p denotes the number of lags of r_t^{SP} and q denotes the number of lags of r_t^{VIX} .

A grid search over $p, q \in 0, \dots, 10$ using the Bayesian Information Criterion (BIC) selects $p^* = 1$ and $q^* = 7$ as the optimal lag orders. Hence, an ADL(1,7) model is estimated.

Coefficients:						
	Estimate	Std. Error	t value	Pr(> t)		
(Intercept)	0.0361384	0.0085485	4.227	2.39e-05 ***		
L(rsp, 1)	-0.1011497	0.0107749	-9.388	< 2e-16 ***		
L(rvix, 0)	-0.1232646	0.0012670	-97.286	< 2e-16 ***		
L(rvix, 1)	-0.0179012	0.0018365	-9.748	< 2e-16 ***		
L(rvix, 2)	-0.0080017	0.0012734	-6.284	3.47e-10 ***		
L(rvix, 3)	-0.0068904	0.0012758	-5.481	6.81e-08 ***		
L(rvix, 4)	-0.0044815	0.0012748	-3.515	0.000441 ***		
L(rvix, 5)	-0.0005877	0.0012729	-0.462	0.644326		
L(rvix, 6)	-0.0027529	0.0012719	-2.164	0.030460 *		
L(rvix, 7)	-0.0074816	0.0012680	-5.900	3.77e-09 ***		

Signif. codes:	0 ****	0.001 ***	0.01 **	0.05 *	0.1 .	1
Residual standard error:	0.7878	on 8498 degrees of freedom				
Multiple R-squared:	0.5319,	Adjusted R-squared:	0.5314			
F-statistic:	1073	on 9 and 8498 DF,	p-value:	< 2.2e-16		

t test of coefficients:						
	Estimate	Std. Error	t value	Pr(> t)		
(Intercept)	0.03613837	0.00810237	4.4602	8.293e-06 ***		
L(rsp, 1)	-0.10114974	0.02551909	-3.9637	7.440e-05 ***		
L(rvix, 0)	-0.12326458	0.00497228	-24.7904	< 2.2e-16 ***		
L(rvix, 1)	-0.01790122	0.00246555	-7.2605	4.195e-13 ***		
L(rvix, 2)	-0.00800169	0.00229106	-3.4926	0.0004808 ***		
L(rvix, 3)	-0.00689038	0.00178439	-3.8615	0.0001135 ***		
L(rvix, 4)	-0.00448151	0.00151020	-2.9675	0.0030108 **		
L(rvix, 5)	-0.00058768	0.00192027	-0.3060	0.7595824		
L(rvix, 6)	-0.00275290	0.00141745	-1.9421	0.0521521 .		
L(rvix, 7)	-0.00748160	0.00229732	-3.2567	0.0011317 **		

Signif. codes:	0 ****	0.001 ***	0.01 **	0.05 *	0.1 .	1

(a) Summary of Model

(b) HAC-robust standard errors

Figure 55: ADL(1,7) model

Interpretation Based on the HAC-robust inference (Figure 55), the autoregressive coefficient $\hat{\beta}_1 \approx -0.101$ is statistically significant, indicating short horizon return reversal where a higher S&P 500 return yesterday is associated with a lower return today, while holding the VIX terms fixed. Economically, this is consistent with short run mean reversion effect and rapid price correction in daily equity returns.

Turning to the VIX terms, the contemporaneous VIX return r_t^{VIX} has a large and highly significant negative coefficient. This means that a 1% increase in today's VIX return is associated with about a 0.123% decrease in the S&P 500 return on the same day, while holding lagged S&P 500 returns and lagged VIX returns fixed. This captures the standard "risk off" co movement, where days with sharp increases in implied volatility tend to coincide with declines in equity prices.

The distributed lag coefficients $\hat{\delta}_1, \dots, \hat{\delta}_7$ are also mostly negative, with several remaining statistically significant under HAC (notably lags 1–4 and 7, while lag 5 is not significant and lag 6 is marginal). This suggests that increases in volatility are not only linked to immediate equity declines (through r_t^{VIX}), but that periods of rising volatility coincide with continued weak equity performance for several subsequent trading days. The fact that more recent lags (in particular r_{t-1}^{VIX} to r_{t-4}^{VIX}) are statistically significant indicates that this effect is concentrated over short horizons, while the insignificance of r_{t-5}^{VIX} and only marginal significance of r_{t-6}^{VIX} suggest the incremental impact fades as the volatility movement becomes less recent. The significance of r_{t-7}^{VIX} indicates some persistence at a one week horizon, consistent with volatility conditions affecting risk appetite over roughly a trading week timescale.

Summing the coefficients on r_t^{VIX} through r_{t-7}^{VIX} yields a total distributed lag effect of approximately -0.17 . This implies that increases in VIX returns are associated with lower S&P 500 returns not only on the same day, but also through the subsequent trading days, indicating that the volatility return relationship is persistent rather than fully absorbed within a single day.

Lastly, the ADL(1,7) model attains an adjusted R^2 of about 0.53, which indicates that 53% of the variation in daily S&P 500 returns is explained by lagged S&P 500 returns together with current and lagged VIX returns. Relative to the AR(1) benchmark, this higher adjusted R^2 suggests that including VIX returns substantially improves the model's in sample explanatory power.

13.5 Instrumental Variable Estimation

In this section, we explore the use of a suitable IV candidate, as returns of VIX is likely an endogenous variable. This is done in iv.R.

$$r_t^{\text{SP}} = \alpha + \beta r_t^{\text{VIX}} + u_t,$$

$$\text{Cov}(r_t^{\text{VIX}}, u_t) \neq 0.$$

This means that VIX returns are correlated with the error term. For example, when there is adverse news such as the outbreak of a war, investors rush to buy put options and volatility protection. This demand pushes implied volatility (and hence r_t^{VIX}) up, while at the same time the S&P 500 return r_t^{SP} falls. Thus, the same underlying shock drives both r_t^{VIX} and the error term u_t , violating the exogeneity assumption.

For a variable to be a suitable IV candidate, it has to satisfy two conditions. First, it must be *relevant*: the instrument Z_t must be correlated with the endogenous regressor r_t^{VIX} , i.e.

$$\text{Cov}(Z_t, r_t^{\text{VIX}}) \neq 0,$$

so that variation in Z_t induces meaningful variation in r_t^{VIX} . Second, it must be *exogenous*: the instrument must not be correlated with the structural error term u_t in the equation,

$$\text{Cov}(Z_t, u_t) = 0,$$

and it should affect r_t^{SP} only through its impact on r_t^{VIX} , not through any other direct channel. We acknowledge that it is very difficult to look for IV candidates that perfectly satisfies these 2 conditions, but we explore the use of 2 potential IV candidates.

13.5.1 Economic Policy Uncertainty Index

The US Daily Economic Policy Uncertainty (EPU) Index (Baker et al., 2016) is a news-based measure of how much US newspapers talk about “economic policy uncertainty.” Each day, it counts articles that simultaneously mention the economy, uncertainty, and at least one policy related term such as legislation, deficit, regulation, congress, the Federal Reserve, or the White House. This raw count is then normalised by the total number of articles to correct for the growing number of newspapers over time. Intuitively, an increase in policy uncertainty mentioned in the news will cause the EPU index to spike.

We explore the use of the EPU index Z_t as a potential IV for VIX returns r_t^{VIX} . With respect to the first IV condition (relevance), spikes in policy uncertainty should increase investors’ demand for protection and raise implied volatility, so we expect

$$\text{Cov}(Z_t, r_t^{\text{VIX}}) \neq 0.$$

This means that on days with higher EPU should be associated with larger movements in VIX returns. For the second condition (exogeneity), we need to assume that, conditional on r_t^{VIX} , shocks to policy uncertainty do not directly affect S&P 500 returns except through their impact

on VIX returns.

$$\text{Cov}(Z_t, u_t) = 0.$$

The EPU index moves r_t^{VIX} but is uncorrelated with the error term u_t , making it a possible (though empirically debatable) IV candidate.

Summary Statistics We work with a transformed version of the EPU index rather than its raw level. Let EPU_t denote the level of the index on day t . We define the daily percentage change in EPU as the scaled log difference

$$\Delta\text{EPU}_t \equiv 100(\log(\text{EPU}_t) - \log(\text{EPU}_{t-1})).$$

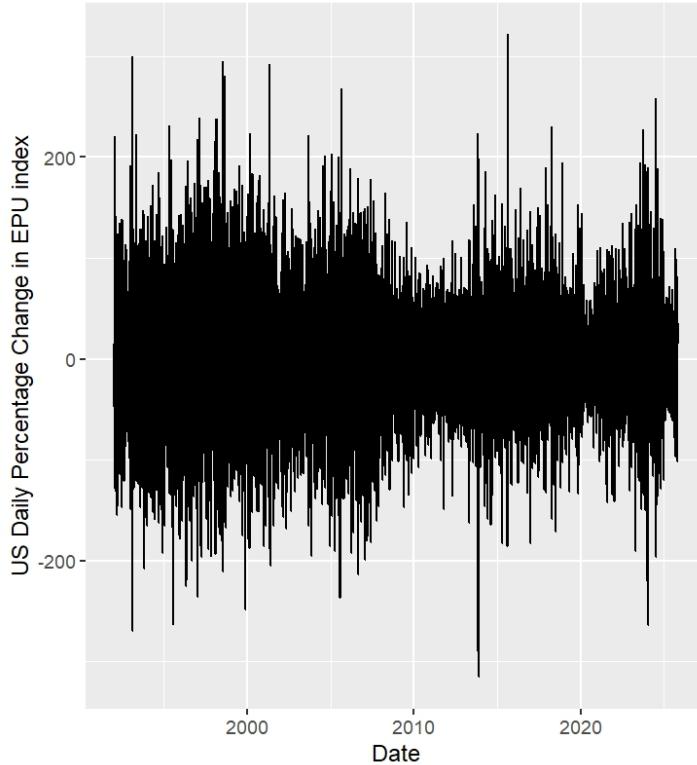


Figure 56: Percentage Change in Daily EPU Index

Table 8: Summary statistics for daily percentage changes in EPU

Mean	Std. dev.	Variance	Min	Max	Skewness	Kurtosis
0.008	61.9	3829	-315	322	0.08	4.2

Results We used the `ivreg()` function in the `AES` package in R, and ran the IV regression using change in EPU index as an IV for returns on VIX. We also used the HAC robust standard errors with 5 lags that is arbitrarily determined.

```

t test of coefficients:

            Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.032880  0.011316  2.9057 0.003674 **
rvix        0.014106  0.090627  0.1557 0.876312
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

```

Figure 57: Coefficients and HAC Robust SE with change in EPU as IV

Interpretation From Figure 57 we can see that the coefficients on VIX returns is positive, which means that if VIX returns increases by 1 unit, returns of S&P 500 increases by 0.014, which contradicts with what we understand their relationship as.

Furthermore, it is statistically insignificant with a large p value. This means that the change in EPU index may not a suitable IV.

This may be due to the fact that change in EPU does not strongly predict VIX returns in the first stage regression.

```

Call:
lm(formula = rvix ~ epu, data = df_epu)

Residuals:
    Min      1Q  Median      3Q     Max 
-44.252 -3.825 -0.402   3.254  76.642 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) -0.002034  0.073722 -0.028   0.9780    
epu         0.002129  0.001191   1.787   0.0739    
---

```

Figure 58: Regression of VIX returns on Change in EPU

Although the change in EPU index could seem like a suitable IV initially, in actual fact, the data disagrees. Conceptually, the EPU index is a weak instrument for daily VIX returns: its day to day changes are only weakly related to VIX returns, and, more importantly, policy uncertainty news is likely to affect S&P 500 returns directly rather than only through VIX returns.

13.5.2 Lag of rvix

The next IV candidate that we explore is the first lag of returns of VIX as introduced in (Zhang et al., 2021). In their setting, the authors address the measurement error and simultaneity problems inherent in using implied volatility by employing lagged implied volatility measures as instruments.

The rationale is that implied volatility is persistent, so its lag provides strong predictive power for the current value (satisfying the relevance condition), while the lagged value is predetermined with respect to the current return shock and therefore should be uncorrelated with the error term (satisfying the exogeneity condition).

Results From Figure 59 we can see that the coefficient on VIX returns when using first lag of VIX returns are -0.073 and is statistically significant at 5%.

```

t test of coefficients:

            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.0327052  0.0085691  3.8167 0.0001362 ***
rvix        -0.0725733  0.0315377 -2.3012 0.0214067 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Figure 59: Coefficients and HAC Robust SE with lag of VIX returns as IV

Interpretation This means that if VIX returns increases by 1 unit, returns on S&P 500 decreases by -0.073. This is in line with the “fear gauge” when implied volatility jumps (VIX up), equity prices tend to fall. Lagged VIX is likely a much stronger IV candidate when compared to change in EPU.

However, validity of lag of VIX returns still relies on the assumption that the lag of VIX returns only affects return on S&P 500 through VIX returns.

Lagged VIX returns are unlikely to be a valid instrument because past changes in implied volatility can plausibly affect current S&P 500 returns directly, not only through their impact on current VIX. A large increase in VIX yesterday typically signals a period of heightened uncertainty and risk aversion that may persist, leading investors to reduce equity exposure today. In addition, many trading strategies adjust positions based on past volatility (for example, cutting stock holdings after a volatility spike), which creates a direct link from yesterday’s VIX move to today’s equity returns. For these reasons, lag of VIX returns is likely correlated with the structural error term, making it a unsuitable IV candidate.

14 Conclusion for Regression Analysis

Across the range of linear, non linear, and dynamic models, the central empirical result is that daily S&P 500 returns is negatively related with daily VIX returns. This “fear gauge” relationship remains statistically significant. In the simplest contemporaneous regression, the estimated slope is economically meaningful (about -0.12), and the model fit is high for daily return ($R^2 \approx 0.52$). A key interpretation is that VIX changes capture shifts in risk appetite and demand for volatility protection that occur on the same days as equity prices shifts.

Allowing for non linearity allows us to capture the possible non linear relationship between S&P 500 returns and VIX returns. The quadratic specification yields a positive and statistically significant squared term, implying a diminishing marginal impact. Increases in VIX returns continue to be associated with lower S&P 500 returns, but the incremental effect becomes less negative at very high VIX returns. This pattern is consistent with a partial “saturation” effect, where once volatility shocks are already extreme, additional increases convey less new information. Interaction models mostly reinforce the conclusion that magnitude matters more than sign. That means that conditioning on the sign of VIX returns or on whether yesterday’s S&P 500 return was positive does not produce any strong evidence of a possible relationship. The main exception arises in the tail based interaction: when VIX returns are extremely high (above the 95th percentile), the negative relationship weakens. A possible economic explanation is contrarian risk bearing (“buy the dip”) when fear spikes are unusually large.

The AR(1) suggests mild short horizon return reversal (a negative and significant lag coefficient), but with negligible explanatory power (adjusted R^2 close to zero). Once current and lagged VIX returns are included in an ADL(1,7) model, the adjusted R^2 increases sharply (adjusted $R^2 \approx 0.53$), and the relationship displays persistence. Several VIX lags remain negative and significant, with effects concentrated over roughly one trading week. Across the contemporaneous and lag coefficients, the total distributed lag impact is around -0.17 , indicating that volatility shocks coincide with continued weakness in equity returns beyond the impact day, rather than being fully absorbed immediately.

Finally, the instrumental variable candidates we tested, the EPU index of lag of VIX returns. Using changes in the Economic Policy Uncertainty (EPU) index produces an unstable and statistically insignificant IV estimate with an unexpected sign. This suggests that EPU is likely a weak IV and the possibility that policy uncertainty affects equity returns directly (violating exclusion). Using lagged VIX returns yields a negative and significant IV estimate, but its validity is also debatable because lagged volatility movements are likely to influence today's equity returns through channels other than today's VIX return.

References

- Allen, D. E., McAleer, M., Powell, R., and Singh, A. K. (2013). A non-parametric and entropy based analysis of the relationship between the vix and s&p 500. *Journal of Risk and Financial Management*, 6(1):6–30.
- Baker, S. R., Bloom, N., and Davis, S. J. (2016). U.s. economic policy uncertainty index data.
- Rašiová, B. and Árendáš, P. (2022). Copula approach to market volatility and technology stocks dependence. Ssrn working paper, SSRN. Preprint submitted to Elsevier, September 30, 2022.
- Sun, Y. and Wu, X. (2009). A nonparametric study of dependence between sp 500 index and market volatility index (vix). *Unpublished manuscript*. Working paper, University of Guelph and Texas AM University, October 28, 2009.
- Wu, X. (2007). Exponential series estimator of multivariate densities. Technical report, Texas A&M University. Working paper / SSRN.
- Zhang, W., Sun, T., Ma, Y., and Wang, Z. (2021). New evidence on the information content of implied volatility of s&p 500: Model-free versus model-based. *Romanian Journal of Economic Forecasting*.