PROJECT 2

Mengzhen Shang 140028
Liyuan Cao 140029

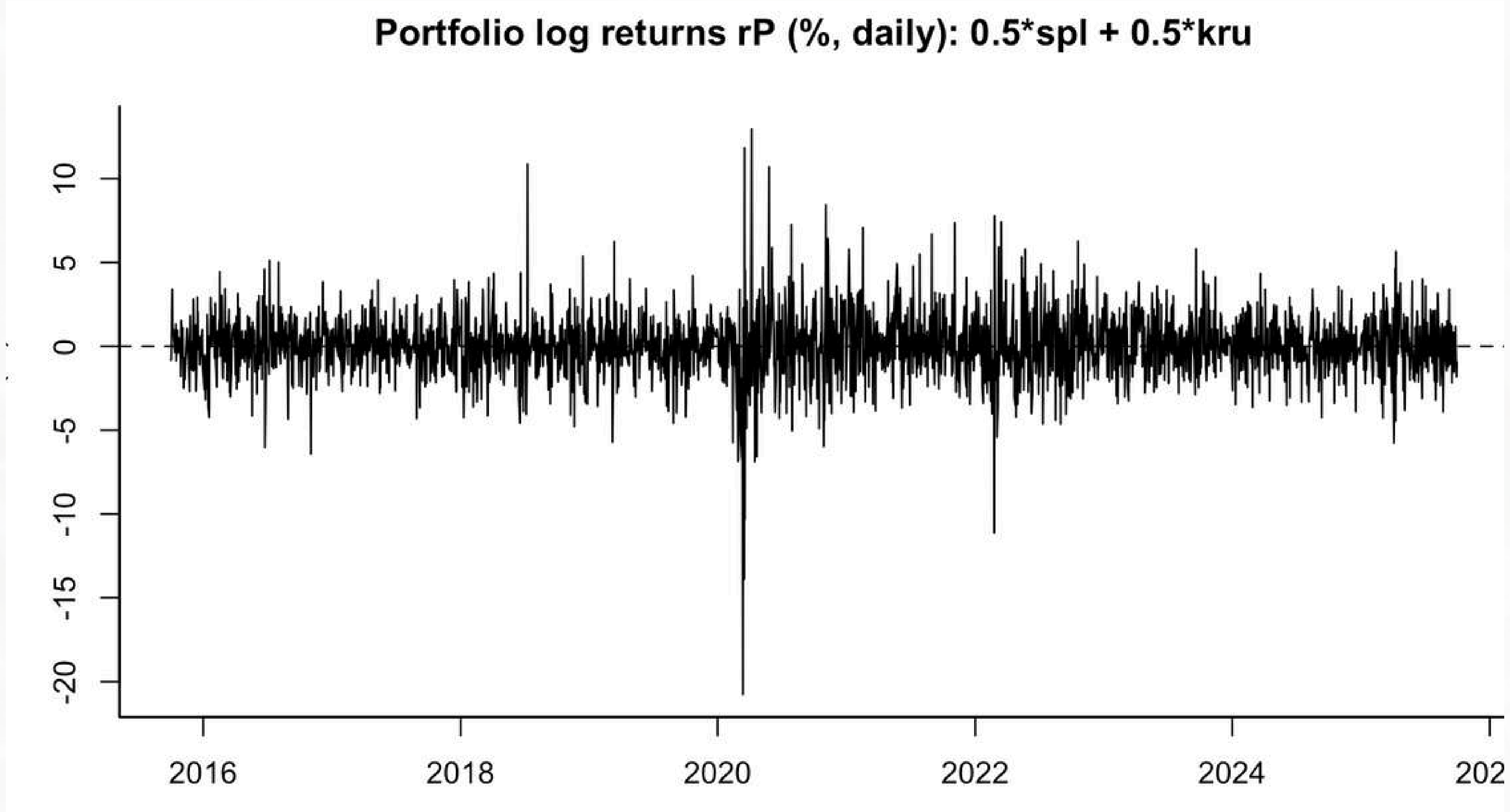
a.Portfolio Returns Overview

Selected 2 assets: spl and kru(weight: 50/50)

SPL (Santander Bank Polska), a major Polish bank offering full-service financial products, and KRU (KRUK S.A.), a European debt management firm.

- time series plot
- moments
- QQ plot
- ACF, ACF of squares returns

a.Portfolio Returns Overview-time series plot



The figure shows daily portfolio returns moving around zero, which means there is no clear trend in the average return.

Volatility changes over time, with quiet periods followed by periods of large movements, especially around 2020.

Some very large returns appear, suggesting heavy tails and changing risk over time.

a.Portfolio Returns Overview-moments

Table: (a) Portfolio moments

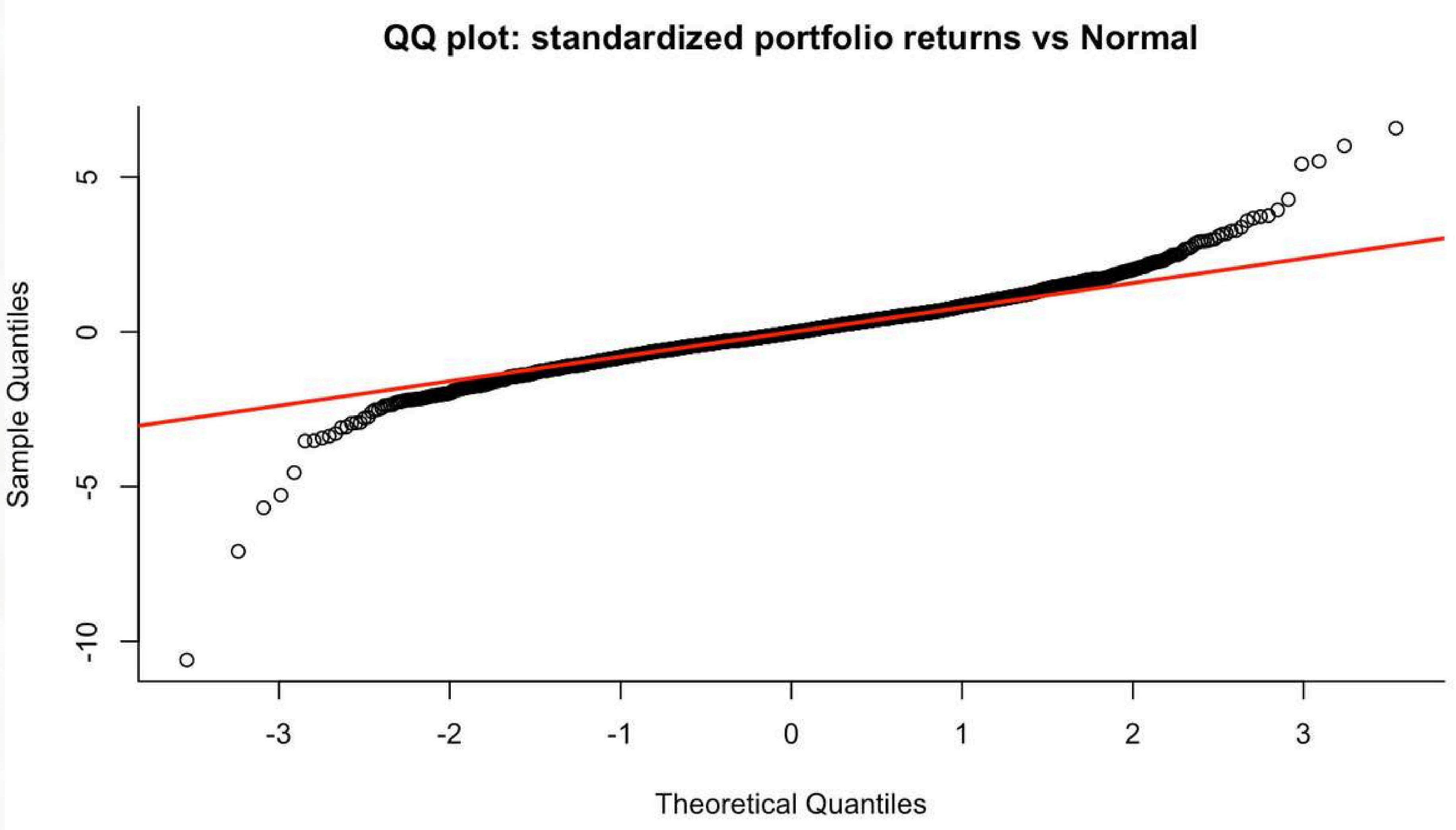
	value
Nyear	249.663
mu (annual)	10.382
sigma (annual)	31.029
min (daily)	-20.771
max (daily)	12.959
skewness	-0.294
kurtosis	12.141
JB stat	8736.547
JB p-value	0.000

- Both daily and annualized volatility are relatively high. Returns are slightly negatively skewed and have very high kurtosis, which means the distribution is not normal and has heavy tails. This is confirmed by the Jarque–Bera test, which clearly rejects normality.

- *Note:Jarque–Bera test: checks whether returns are normally distributed (small p-value → normality rejected).*

  H_0: Returns follow a normal distribution.
  H_1: Returns do not follow a normal distribution.

a.Portfolio Returns Overview-QQ plot

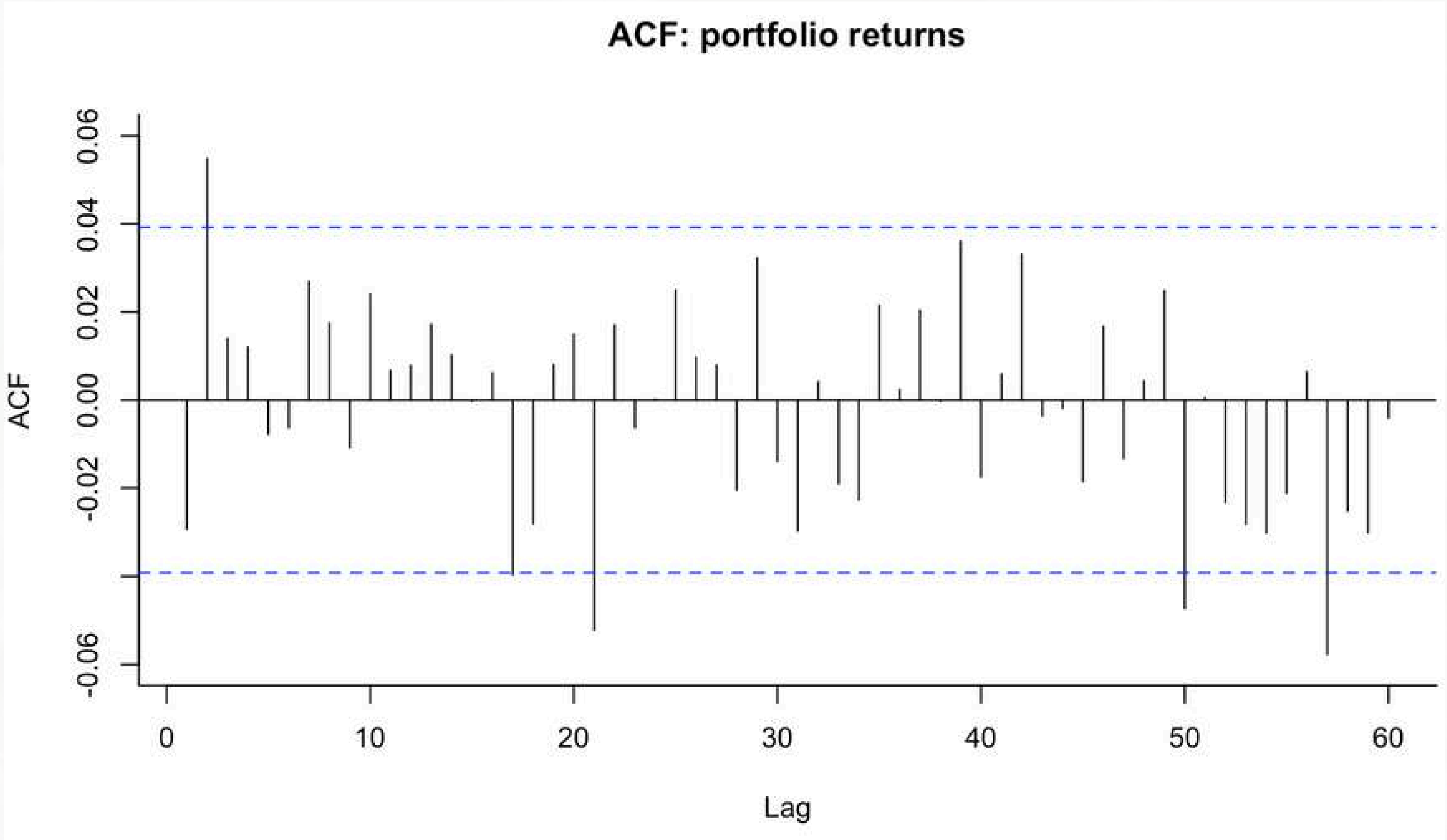


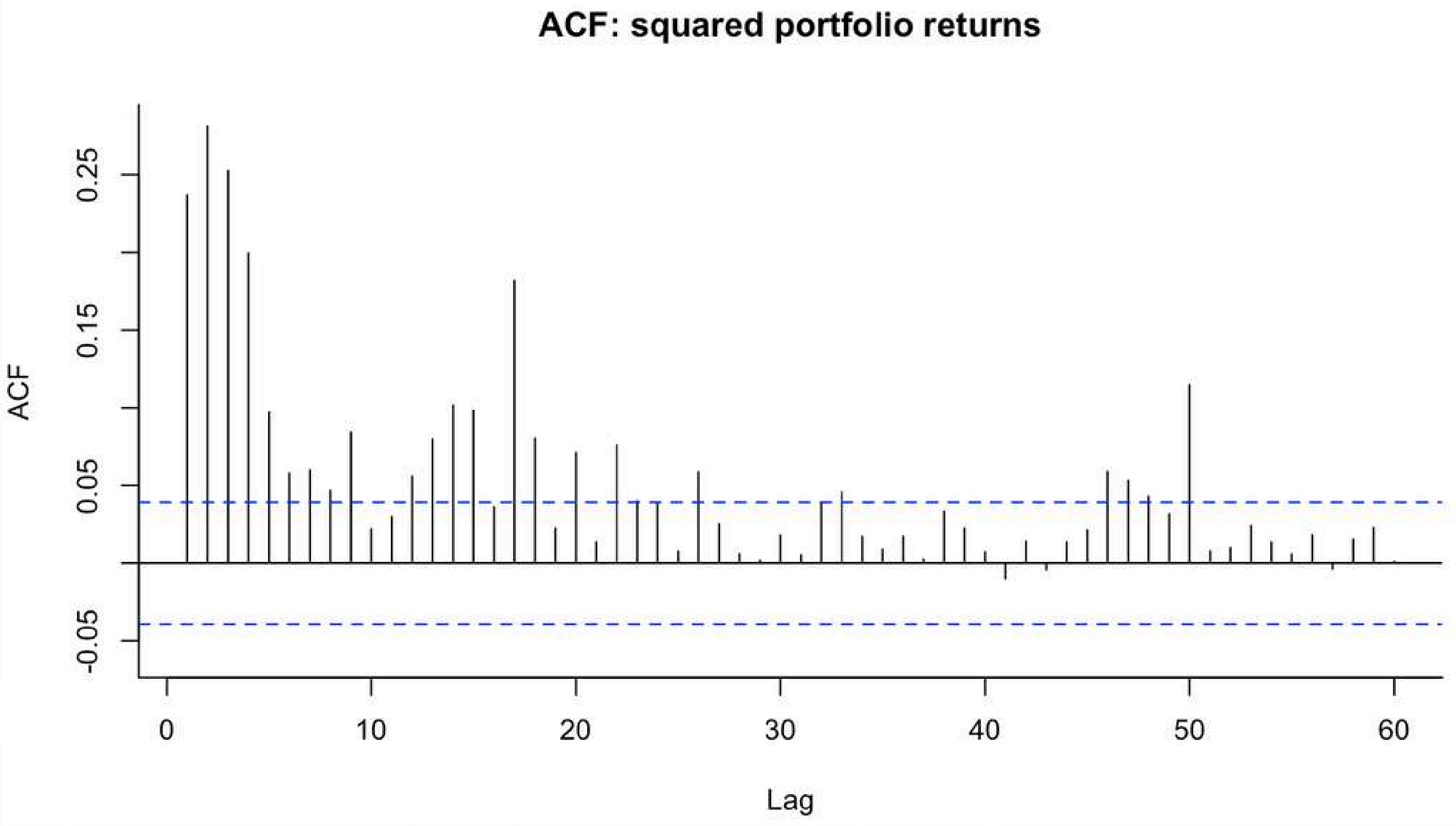
The QQ plot compares standardized portfolio returns with a normal distribution.

While observations in the center follow the reference line reasonably well, strong deviations appear in both tails.

This indicates that extreme returns occur more often than under normality.

a.Portfolio Returns Overview-ACF, ACF of squares returns.





- The ACF plot shows that most autocorrelation coefficients are small and lie within the confidence bounds.
- This suggests that portfolio returns exhibit little serial correlation and are close to white noise.

- The ACF of squared portfolio returns shows many significant autocorrelations over multiple lags.
- This indicates volatility clustering, meaning that periods of high volatility tend to be followed by high volatility, and periods of low volatility by low volatility.

b. Univariate GARCH Modeling

In Part a， returns are heavy-tailed and show volatility clustering, so more flexible models such as GARCH are required.

Best model: eGARCH(1,1) with Student-t errors

- Model selection
- Parameter estimates
- Conditional tandard deviation
- Diagnostic check

b. Univariate GARCH Modeling-Model selection (AIC / BIC)

```
=== Model selection table (sorted by BIC) ===
> print(knitr::kable(sel_BIC, digits = 4))


|   |Model    |Dist |        LL|    AIC|    BIC|
|:--|:--------|:----|---------:|------:|------:|
|8  |eGARCH   |std  | -4958.846| 3.9735| 3.9874|
|11 |gjrGARCH |std  | -4963.699| 3.9774| 3.9913|
|9  |eGARCH   |ged  | -4965.343| 3.9787| 3.9926|
|12 |gjrGARCH |ged  | -4968.848| 3.9815| 3.9955|
|2  |sGARCH   |std  | -4973.289| 3.9842| 3.9959|
|5  |iGARCH   |std  | -4984.333| 3.9923| 4.0016|
|3  |sGARCH   |ged  | -4980.750| 3.9902| 4.0018|
|6  |iGARCH   |ged  | -4992.205| 3.9986| 4.0079|
|7  |eGARCH   |norm | -5008.945| 4.0128| 4.0244|
|10 |gjrGARCH |norm | -5010.593| 4.0141| 4.0257|
|1  |sGARCH   |norm | -5030.042| 4.0288| 4.0382|
|4  |iGARCH   |norm | -5043.627| 4.0389| 4.0459|
```

We compare several univariate GARCH specifications using log-likelihood and information criteria.

Models are compared using AIC and BIC. The best model is eGARCH(1,1) with Student-t errors, as it has the lowest AIC and BIC in the table.

The preference for the Student-t distribution is consistent with the heavy tails observed in the data, while the eGARCH structure provides a flexible specification for time-varying volatility.

```
Best model by BIC:
> cat(sprintf("  %s(1,1) with %s distribution\n", best_row$Model, best_row$Dist))
  eGARCH(1,1) with std distribution
```

b. Univariate GARCH Modeling-Parameter estimates

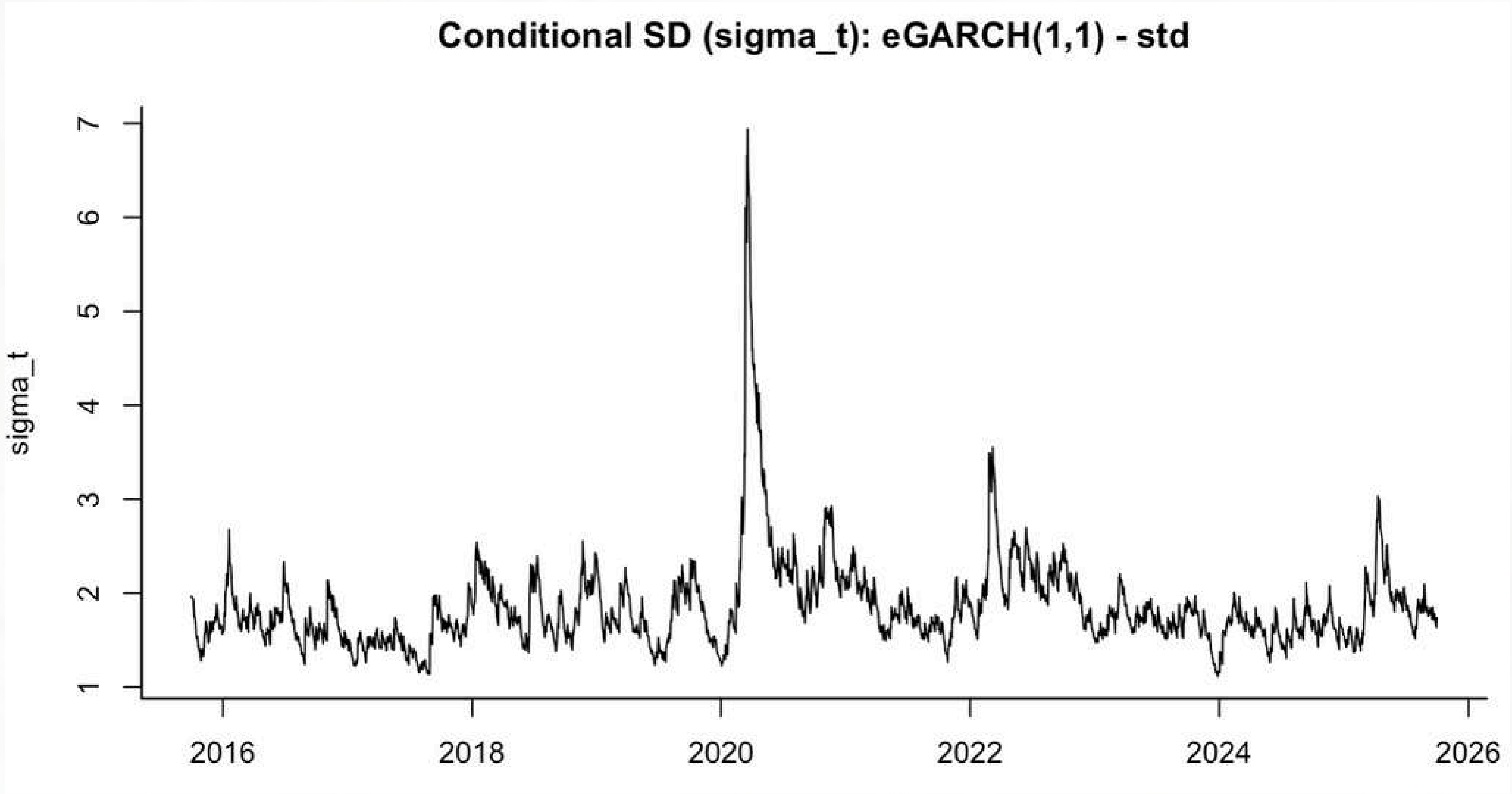
```
=== Parameter estimates (full table) ===
> print(knitr::kable(param_tbl, digits = 6))


|Parameter |   Estimate|  StdError|     tValue|   pValue|
|:---------|----------:|---------:|----------:|--------:|
|mu        |   0.012642|  0.032120|   0.393577| 0.693894|
|omega     |   0.028990|  0.002991|   9.693277| 0.000000|
|alpha1    |  -0.064549|  0.011925|  -5.412782| 0.000000|
|beta1     |   0.976132|  0.001421| 686.831804| 0.000000|
|gamma1    |   0.117263|  0.023779|   4.931323| 0.000001|
|shape     |   6.777835|  0.894486|   7.577348| 0.000000|
```

- The table reports parameter estimates for the selected **eGARCH(1,1) model with Student-t errors**.
- The mean parameter(mu) is not significant, but the volatility parameters are highly significant. Volatility persistence is strong since beta1 is close to one, and the asymmetry term(gamma1) is significant, indicating that volatility responds differently to shocks of different signs.
- The Student-t shape parameter is also significant, confirming heavy tails in returns.

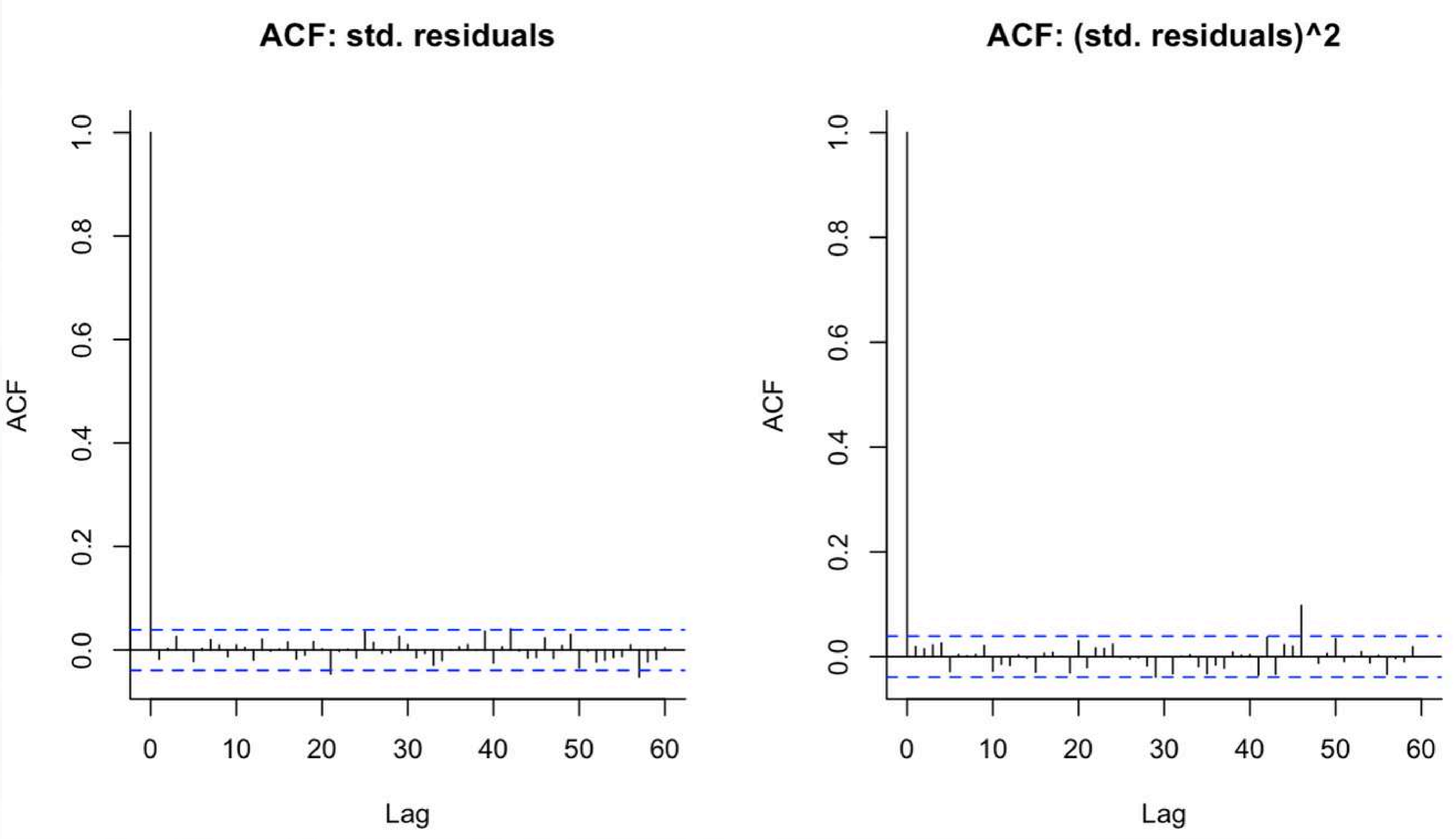
b. Univariate GARCH Modeling-Conditional tandard deviation (plot)

- This figure shows the conditional standard deviation (σ_t) estimated from the selected eGARCH(1,1) model with Student-t errors.
- Volatility is clearly time-varying and exhibits strong clustering.
- A sharp increase is observed around 2020, corresponding to a period of market stress, followed by a gradual decline.
- This confirms that the model captures changing risk over time.



b. Univariate GARCH Modeling- Diagnostic check

- These ACF plots are model diagnostics.
- The autocorrelations of standardized residuals are mostly within the confidence bands, indicating little remaining serial dependence.
- The ACF of squared standardized residuals is also close to zero, suggesting that the eGARCH model has largely captured the volatility clustering in the data.



c. Copula Estimation and Selection

Part (b) fits GARCH-t models to capture time-varying volatility. So part (c) focuses on the dependence structure between the two assets.

- LL comparison
- Copula estimation

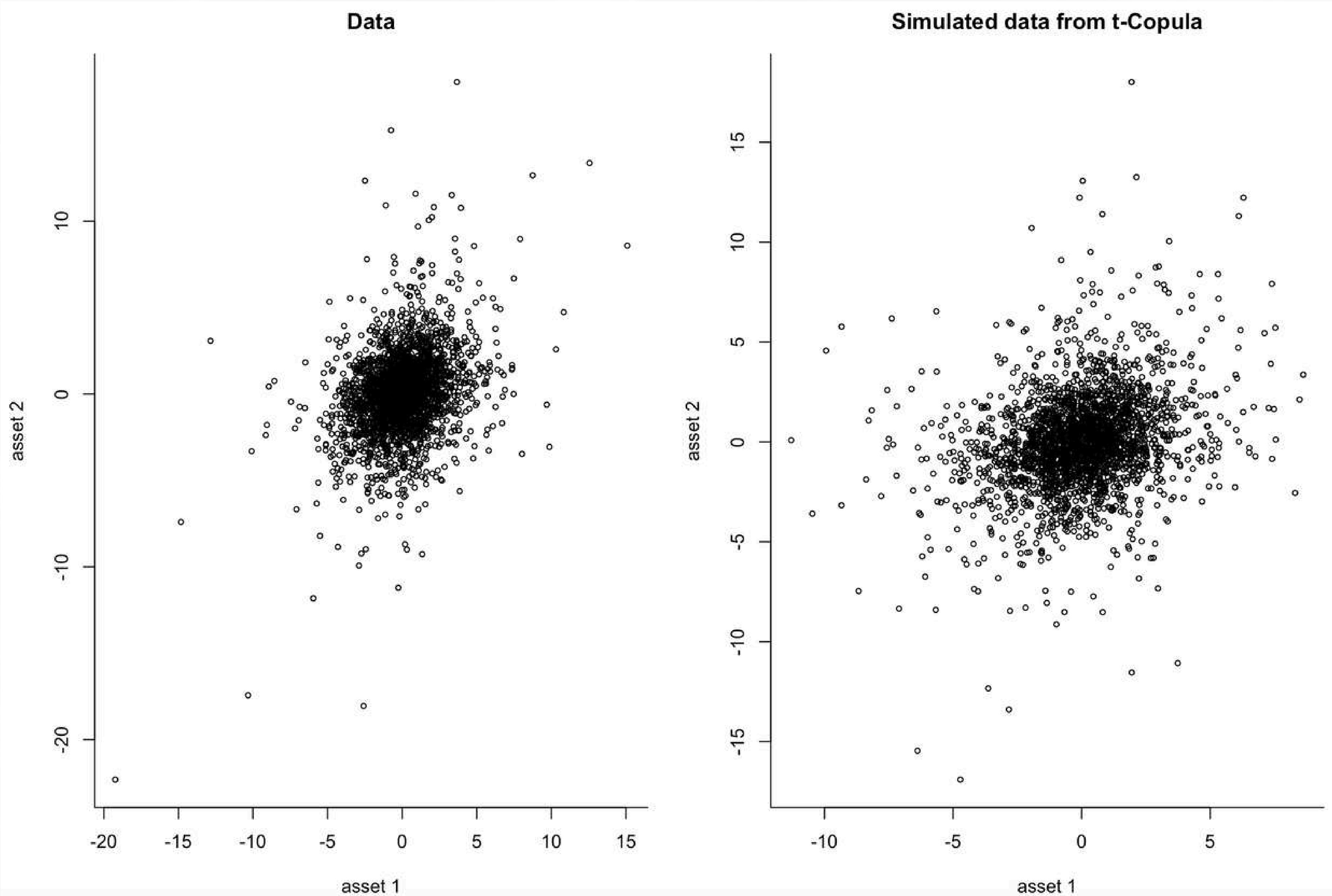
c. Copula Estimation and Selection-LL comparison

```
=== Copula comparison (Log-Likelihood, higher is better)
> print(knitr::kable(LLtab, digits = 3))


|         |Copula   |      LL|
|:--------|:--------|-------:|
|tCopula  |tCopula  | 121.263|
|Gaussian |Gaussian | 116.169|
|Frank    |Frank    | 113.100|
|Gumbel   |Gumbel   | 101.929|
|Clayton  |Clayton  |  80.966|
```

- Several copula families are compared using log-likelihood values.
- The table shows that the **t-copula** achieves the **highest** log-likelihood, and is therefore selected as the best copula.
- This suggests that the dependence structure between the two assets is better captured by a copula that allows for tail dependence, rather than others.

c. Copula Estimation and Selection-Copula estimation



- The left panel shows the scatter plot of historical returns of two assets (spl and kru), revealing non-linear dependence and tail co-movements.
- The right panel presents simulated joint returns based on the fitted t-Copula and GARCH-t marginal models.
- The fitted t-Copula, combined with GARCH-t marginals, successfully replicates the observed dependence structure, including extreme tail co-movements, confirming its suitability for modeling joint financial risks.

c. Copula Estimation and Selection-Copula estimation

```
> cat(sprintf("\nKendall tau (real) = %.3f, (sim) = %.3f\n", tau_real, tau_sim))

Kendall tau (real) = 0.198, (sim) = 0.177
```

- Kendall's tau from the empirical data is 0.198, while the value from the t-copula simulation is 0.177.
- The close values indicate that the selected copula reproduces the dependence strength between the two assets reasonably well.

Note:

Kendall's tau is a rank-based measure of dependence defined as the difference between the probabilities of concordant and discordant pairs of observations. In simple terms, it quantifies how often two variables move in the same direction versus opposite directions, using only the ordering of the data.

Kendall's tau is particularly suitable for copula analysis because it depends only on the dependence structure and is invariant to monotonic transformations of the marginals. This means that tau is not affected by the specific marginal distributions, which is exactly what we want when modeling dependence with copulas.

The value $\tau \approx 0.2$ indicates mild to moderate positive dependence: the two assets tend to move in the same direction more often than in opposite directions, but the dependence is not very strong.

d. VaR & ES Estimation Results

Comparison across horizons (H=1, 10) and probabilities (p=1%, 5%)

- Static Methods: Historical, Normal, t-Student
- Dynamic Methods: EWMA, GARCH, DCC-GARCH
- Dependence Modeling: Best-fit Copula
- Methodology: Multi-step Monte Carlo Simulation (for H=10)

d. Risk Estimation: Static Models Analysis(Benchmarking Historical Data against Gaussian & t-Distributions)

1. Normal Distribution Underestimates Tail Risk

- Evidence: At 1-day 1% level, Normal ES (-5.20%) is significantly lower than Historical ES (-7.24%).
- Implication: Gaussian assumptions fail to capture extreme market crashes ("Fat Tails"), leading to dangerous optimism.

2. t-Student Improvements

- Evidence: t-Student ES (-6.89%) aligns much closer to Historical data.
- Reason: The model explicitly accounts for heavy tails (kurtosis > 3), making it a safer static alternative.

3. Horizon Scaling (10-day)

- The underestimation persists over time. Normal 10-day ES (-16.35%) is still ~2.3 percentage points lower than Historical (-18.64%).

Method	Horizon	Prob	VaR	ES
Historical	1-day	1%	-4.4038	-7.2448
Historical	1-day	5%	-2.7871	-4.2057
Historical	10-day	1%	-15.0905	-18.6399
Historical	10-day	5%	-9.4639	-12.9337
Normal	1-day	1%	-4.6067	-5.1965
Normal	1-day	5%	-3.2286	-4.0528
Normal	10-day	1%	-14.0686	-16.3499
Normal	10-day	5%	-9.6062	-12.3422
t	1-day	1%	-5.1264	-6.8924
t	1-day	5%	-3.0032	-4.3887
t	10-day	1%	-15.1828	-18.3288
t	10-day	5%	-9.9863	-13.2276

d. Risk Estimation: Dynamic Models Analysis(Impact of Volatility Clustering & Recent Market Conditions)

Method	Horizon	Prob	VaR	ES
Historical	1-day	1%	-4.4038	-7.2448
Historical	1-day	5%	-2.7871	-4.2057
Historical	10-day	1%	-15.0905	-18.6399
Historical	10-day	5%	-9.4639	-12.9337

Method	Horizon	Prob	VaR	ES
EWMA	1-day	1%	-3.1557	-3.6619
EWMA	1-day	5%	-2.2715	-2.8417
EWMA	10-day	1%	-10.0972	-11.2320
EWMA	10-day	5%	-6.8387	-8.6593
GARCH	1-day	1%	-4.2167	-5.3594
GARCH	1-day	5%	-2.7797	-3.7538
GARCH	10-day	1%	-14.5531	-17.6219
GARCH	10-day	5%	-9.4106	-12.6110

- 1. EWMA Reflects "Calm" Regimes
  - Observation: EWMA estimates are the lowest among all models (e.g., 1-day 1% VaR is only -3.16%).
  - Reason: EWMA places heavy weight (lambda=0.94) on recent observations. Since the recent sample period (2024-2025) exhibits relatively low volatility compared to historical crashes (2008/2020), the model predicts a stable future.
  - Risk: This may lead to underestimation if a sudden regime shift occurs.
- 2. GARCH Captures Mean Reversion
  - Observation: GARCH 10-day ES (-17.62%) is higher than EWMA (-11.23%) and closer to Historical data.
  - Mechanism: Unlike EWMA, the GARCH model accounts for mean reversion (long-run variance). Even if current volatility is low, GARCH anticipates it will eventually revert to a higher long-term average, leading to more prudent long-horizon forecasts.
- 3. Conclusion on Dynamic Models
  - While responsive to market changes, dynamic models are highly sensitive to the specific window of data used for calibration.

d. Risk Estimation: Advanced Models Analysis(Capturing Dynamic Correlations & Tail Dependence (DCC & Copula)

1. DCC-GARCH: The "Perfect Storm" Scenario

- Critical Finding: DCC predicts the highest extreme loss: -26.78% (10-day 1% ES). This is ~50% higher than the Normal model.
- Reason: Static models assume constant correlations. DCC simulates time-varying correlations, capturing scenarios where assets crash together (correlation spikes to 1), destroying diversification benefits.

2. Copula: Consistent Tail Risk

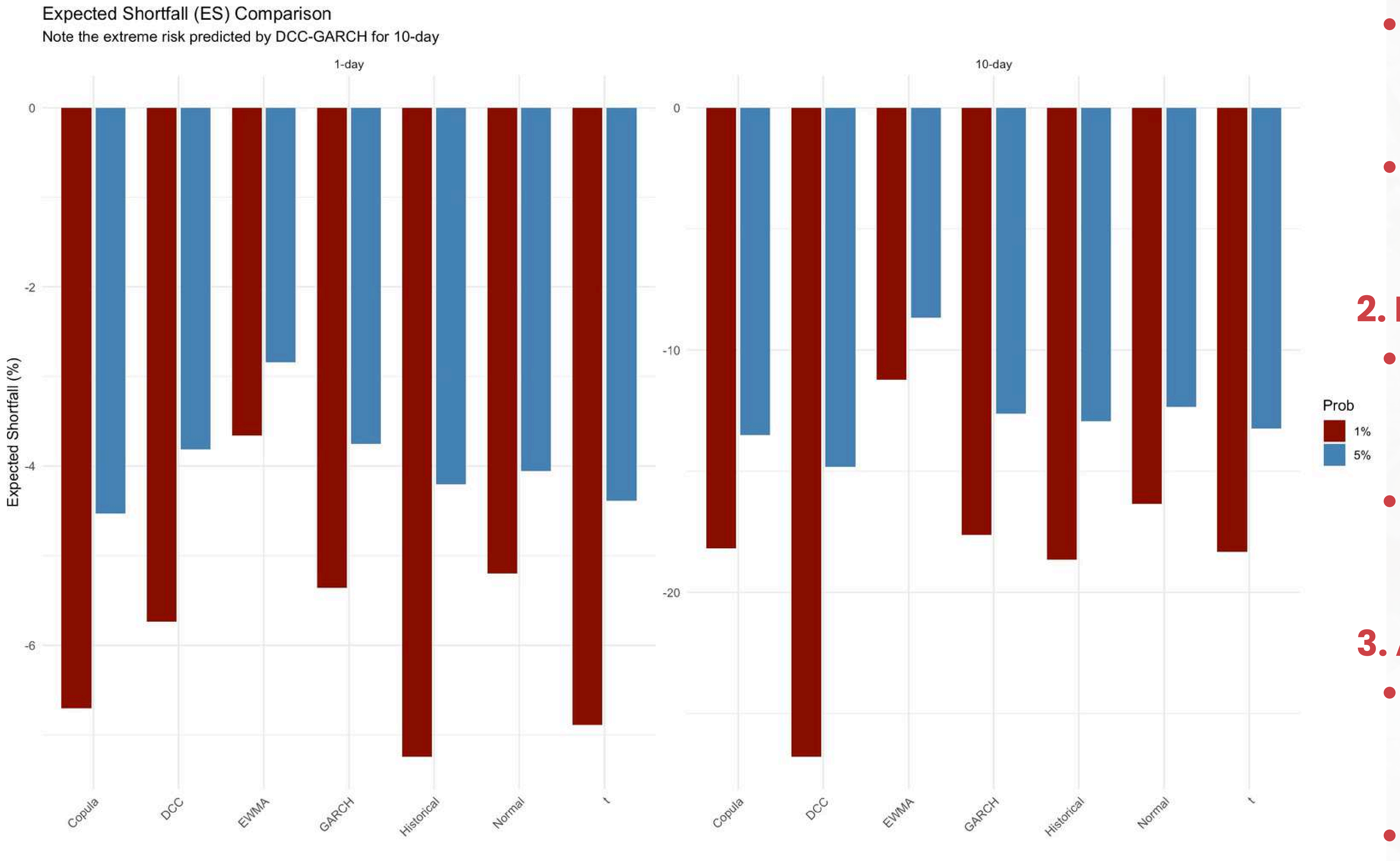
- Observation: The Copula model (likely t-Copula or similar) also predicts high risk (-18.18% for 10-day ES), aligning closely with Historical Simulation.
- Advantage: It mathematically models tail dependence—the likelihood that if one asset crashes, the other will too—better than simple linear correlation.

3. Final Verdict

- For short-term (1-day) stability, most models agree.
- For long-term stress testing (10-day), only DCC reveals the true extent of structural market risk.

Method	Horizon	Prob	VaR	ES
DCC	1-day	1%	-4.4747	-5.7371
DCC	1-day	5%	-2.7186	-3.8122
DCC	10-day	1%	-17.2449	-26.7824
DCC	10-day	5%	-9.1186	-14.8299
Copula	1-day	1%	-5.1672	-6.7034
Copula	1-day	5%	-3.2390	-4.5298
Copula	10-day	1%	-15.3676	-18.1797
Copula	10-day	5%	-10.4142	-13.5003

d. Risk Estimation: Strategic Conclusion(Benchmarking Model Performance under Stress (10-Day Horizon))



1. Static Models: The "Fat Tail" Trap

- Finding: Normal distribution (Gaussian) consistently underestimates risk compared to Historical data.
- Takeaway: Relying on standard deviations is dangerous; markets exhibit heavy tails that only t-Student or Historical methods capture.

2. Dynamic Models: The "Recency" Trap

- Finding: EWMA provides the most optimistic (lowest) risk estimates due to its reliance on recent low-volatility data.
- Takeaway: While responsive, EWMA fails to anticipate regime shifts, making it unsuitable for worst-case stress testing.

3. Advanced Models:The "Correlation" Reality

- Finding: DCC-GARCH predicts extreme losses (-27%) over 10 days, far exceeding all other models.
- Takeaway: It captures the breakdown of diversification (correlations spiking to 1). This confirms that DCC is the superior model for modeling structural market crashes.

e. Model Validation & Backtesting

Evaluating Risk Forecasts for Historical, EWMA, and DCC Models

- Tests Covered: Kupiec, Christoffersen (Ind/CC), Pelletier, Frey-McNeil
- Objective: Testing H0 (Model Correctness) vs H1 (Model Failure)

e. Backtesting Methodology

Statistical Framework for Validating Model Accuracy

1. Kupiec Test (Unconditional Coverage)

- Focus: Frequency.
- Question: Does the number of violations match the expected level (e.g., 1%)?
- Role: A basic check on whether the model is too aggressive or too conservative.

2. Christoffersen Independence (Ind)

- Focus: Clustering.
- Question: Are violations scattered randomly over time?
- Role: Detects if the model fails to adapt to volatility clustering (e.g., failing multiple days in a row).

3. Christoffersen Conditional Coverage (CC)

- Focus: Combined Validity.
- Logic: Mathematically combines Test 1 (Count) + Test 2 (Independence).
- Role: A rigorous "Joint Test". Passing this means the model gets both the count and the timing right.

4. Pelletier Test (Duration)

- Focus: Timing.
- Logic: Examines the time duration between violations using a Weibull distribution.
- Role: A more sensitive alternative to the Independence test for detecting structural flaws.

5. Frey-McNeil Test (for ES)

- Focus: Magnitude.
- Question: When a violation occurs, is the size of the loss predicted correctly?
- Role: The only test that validates Expected Shortfall (tail depth), not just VaR.

e. Critical Assessment of Backtesting Results(Benchmarking Model Robustness: Extreme Risk (1%) vs. Structural Dynamics (5%))

Table: Backtesting: Expected(Exp) vs Actual(Act) Exceedances

Model	Alpha	Exp	Act	Kupiec	Res	Ind	Res	CC	Res	Pell	Res	FM	Res
HS	1%	22.48	36	0.008	Fail	0.020	Fail	0.002	Fail	0.001	Fail	0.252	Pass
HS	5%	112.40	118	0.591	Pass	0.011	Fail	0.033	Fail	0.034	Fail	0.161	Pass
EWMA	1%	22.48	39	0.002	Fail	0.032	Fail	0.001	Fail	0.003	Fail	0.031	Fail
EWMA	5%	112.40	117	0.658	Pass	0.123	Pass	0.276	Pass	0.690	Pass	0.002	Fail
DCC	1%	22.48	6	0.000	Fail	0.858	Pass	0.000	Fail	0.999	Pass	0.517	Pass
DCC	5%	112.40	55	0.000	Fail	0.055	Pass	0.000	Fail	0.053	Pass	0.160	Pass

1. Primary Evaluation: Extreme Tail Risk (1% Level)

- Context: The 1% level represents crisis conditions, serving as the primary stress test for capital adequacy.
- Directional Divergence:
- HS & EWMA (Aggressive Failure): Significantly underestimate risk (Act>>Exp), leaving the portfolio exposed to unexpected losses.
- DCC (Conservative Bias): Overestimates risk (Act<<Exp). While statistically a rejection, this provides a safety buffer in a risk management context.

2. Structural Diagnosis (5% Level)

- Diagnostic Function: Results at the 5% level reveal internal model mechanics.
- The "Clustering" Flaw: Although Historical Simulation (HS) passes the frequency test (Kupiec) at 5%, it fails the Independence Test.
- Implication: This confirms HS cannot adapt to volatility clustering, explaining its collapse under the stricter 1% test.

3. Strategic Conclusion

- Verdict: DCC-GARCH is the preferred model.
- Reason: It is the only model that ensures survival during tail events (1%) and maintains structural integrity (Independence/Pelletier Pass) across all levels.

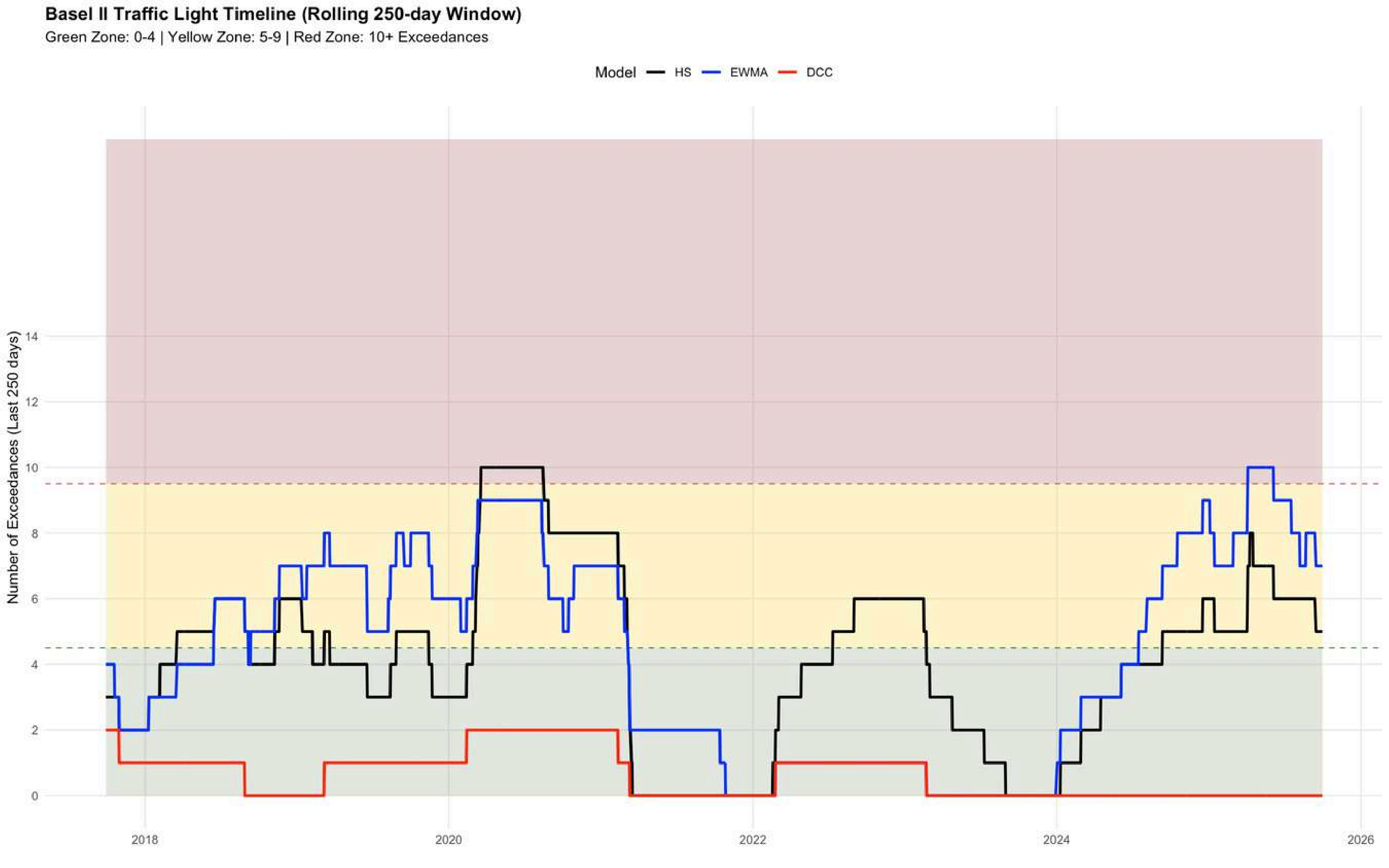
e. Backtesting risk models(Basal Committee “Traffic lights” approach)

Table: Final Regulatory Capital Charge Determination

Model	Exceedances	Zone	Plus_Factor_k	Total_Multiplier	cdf_percent
HS	5	Yellow	0.40	3.40	95.88%
EWMA	7	Yellow	0.65	3.65	99.6%
DCC	0	Green	0.00	3.00	8.11%

Table: Basel Traffic Light Test based on Rolling 250-day Windows

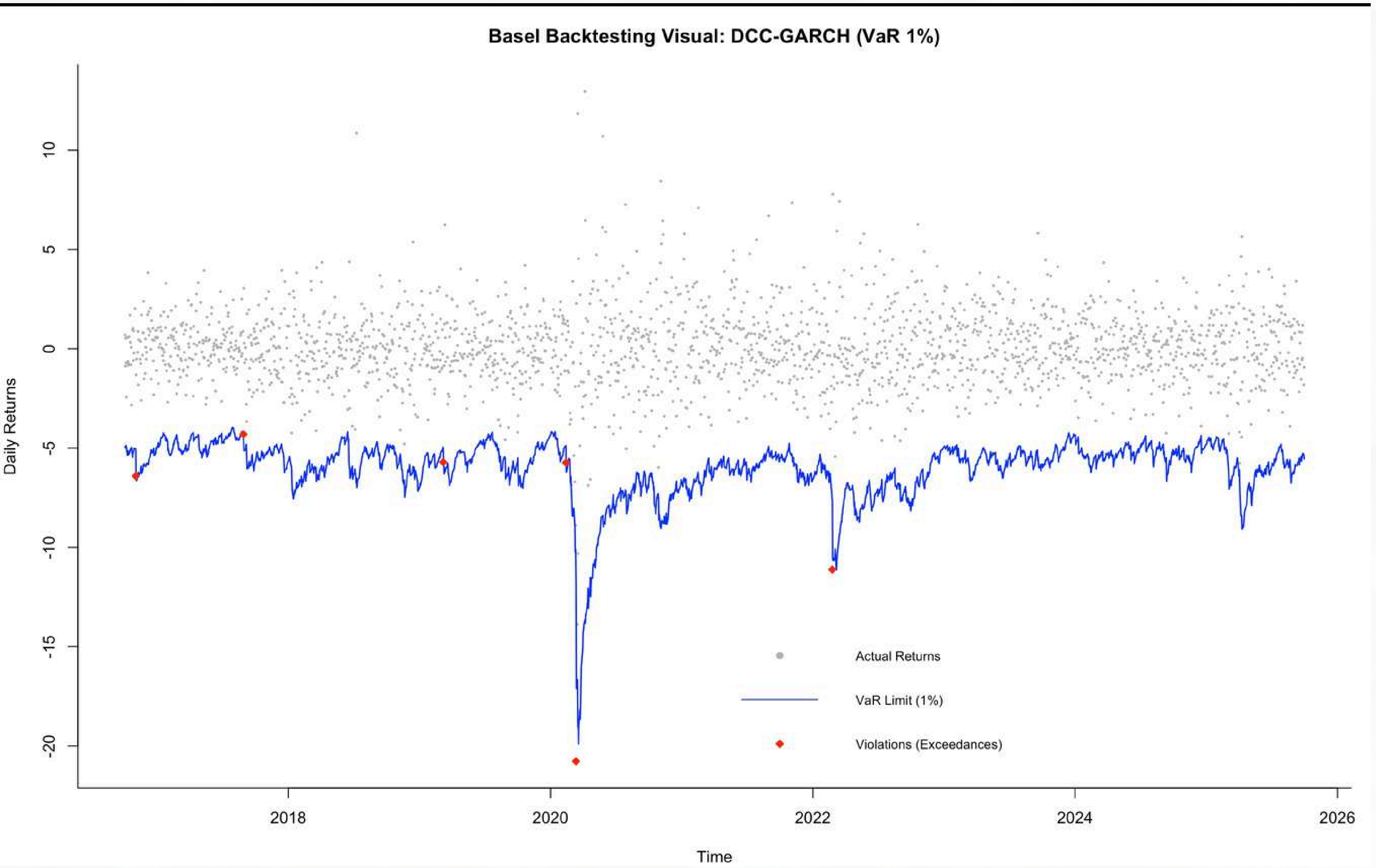
Model	Worst_Window	Green_Pct	Yellow_Pct	Red_Pct
HS	Red	0.543	0.405	0.052
EWMA	Red	0.511	0.469	0.021
DCC	Green	1.000	0.000	0.000



1. Static vs. Dynamic Assessment
   - Current Snapshot (Table 1): Both HS and EWMA fall into the Yellow Zone, incurring capital penalties (Plus Factor k = 0.40 & 0.65).
   - The Trap: A single snapshot masks the true risk. The Rolling Window Test (Table 2) reveals that HS and EWMA are not just "imperfect"—they are dangerous.
2. Critical Failure of Simple Models
   - The "Red Zone" Breach: As shown in the timeline chart, HS and EWMA spike above 10 exceedances during stress periods (e.g., 2020, 2025).
   - Regulatory Consequence: Reaching the Red Zone (>99.99% probability of failure) implies model rejection and maximum penalties.
   - Reason: These models failed to capture the fat-tailed risk inherent in the SPL (Bank) + KRUK (Debt Mgmt) portfolio.
3. Why DCC-GARCH Wins
   - 100% Regulatory Compliance: The DCC model remains in the Green Zone (0-4 exceedances) throughout the entire 8-year history.
   - Robustness: With low degrees of freedom (df≈5.6), DCC accounts for extreme tail risks and dynamic correlations between the banking and debt sectors.
   - Conclusion: While conservative, DCC is the only model that ensures survival during market crises.

e. Visualizing Model Dynamics: DCC-GARCH Backtest (1%)

Evidence of Volatility Adaptation and Independence of Exceedances



1. Extreme Reactivity (The Blue Line)

- Observation: Notice how the VaR Limit (Blue Line) is not smooth. It reacts violently to market shocks, specifically the deep "V-shape" crash during the 2020 COVID-19 crisis.
- Mechanism: This demonstrates Volatility Clustering in action. The GARCH component immediately widened the risk buffer as volatility spiked, effectively "shielding" the portfolio from consecutive failures.

2. The "Clean" Backtest (The Red Diamonds)

- Scarcity: Over ~2,250 trading days, there are only ~6 violations (Red Diamonds). This is visually consistent with the "Green Zone" status shown in the previous slide.
- Independence: Crucially, the red diamonds are scattered, not clustered together. This visually confirms the Christoffersen Independence Test pass—meaning the model adapts fast enough that a failure today doesn't guarantee a failure tomorrow.

3. The "Fat Tail" Buffer

- Why so conservative? You can see a significant gap between the returns (gray dots) and the VaR line during calm periods.
- The Cause: The estimated degrees of freedom (df≈5.6) forces the model to assume "extreme events are always possible," creating a deeper safety margin than Normal Distribution models (EWMA).

e. Backtesting risk models(Conclusion)

1. Historical Simulation (HS): "The Rear-View Mirror"
   - Statistical Failure:
     - Independence Fail: Fails to capture volatility clustering ($p < 0.05$ in Ind/CC tests). It assumes risk is constant, ignoring that "bad days come in clusters."
     - Tail Risk: Significant underestimation of risk (Act 36 > Exp 22.5).
   - Economic & Regulatory Impact:
     - Regulatory Nightmare: Hits the "Red Zone" 5.2% of the time in rolling tests. This implies a high probability of regulatory intervention or model rejection during crises.
     - Assessment: Too slow to react. Dangerous for a dynamic portfolio like SPL+KRUK.
2. EWMA: "The Fair-Weather Friend"
   - Statistical Failure:
     - Distribution Flaw: Relies on the Normal Distribution. It fails the 1% Kupiec test drastically (39 exceedances) because it ignores "Fat Tails" (Kurtosis).
     - Calibration: Fails the Frey-McNeil test, meaning even when it predicts a loss, it underestimates the magnitude of that loss.
   - Economic & Regulatory Impact:
     - Capital Trap: While it appears "efficient" in calm markets, it crashes into the "Red Zone" (2.1%) during shocks.
     - Assessment: Good for daily volatility monitoring, but unfit for regulatory capital purposes due to tail blindness.
3. DCC-GARCH: "The Iron Shield"
   - Statistical Strength (with a Caveat):
     - Robustness: Passes all Independence tests (Ind, Pell). It perfectly adapts to changing correlations between Banking (SPL) and Debt Management (KRUK).
     - The "Conservative" Fail: Fails Kupiec (1%) only because it is too safe (6 exceedances vs 22.5 expected).
   - Economic & Regulatory Impact:
     - Survival: 100% Green Zone in rolling history. Zero regulatory fines (k=0).
     - Trade-off: It sacrifices Capital Efficiency (higher reserves) for Absolute Solvency.
     - Assessment: The only model that accounts for the fat-tailed, non-linear nature of the portfolio.

Conclusion: In financial risk management, survival outweighs efficiency. While HS and EWMA offer lower capital charges during calm periods, their failure to withstand stress tests (Red Zone) makes them non-viable. DCC-GARCH, despite its conservatism, is the only compliant and robust choice for the SPL/KRUK portfolio.

Thank You~