Copula modelling for stock returns

THE UNIVERSITY OF MELBOURNE

Student: Hespera Henzell   Supervisor: Dr Pavel Krupskiy

THE UNIVERSITY OF MELBOURNE

What are Copulas?

According to Sklar's theorem [1], for every joint distribution function $\boldsymbol{F}_{\{1:d\}}$ of continuous random variables $X_1, \dots, X_d$ with margins $F_{X_1}, \dots, F_{X_d}$, there exists a unique copula C such that:

$$\boldsymbol{F}_{\{1:d\}}(x_1, \dots, x_d) = C\left(u_1, \dots, u_d\right), \quad u_i = F_{X_i}(x_i)$$

- C is acting as a distribution function (DF) of the margins
- Modelling margins and the dependence separately is often easier
- We can even change the distribution of the margins in the copula [2]
- We can obtain a copula DF by inverting Sklar's theorem:

$$C\left(u_1, \dots, u_d\right) = \boldsymbol{F}_{\{1:d\}}\left(F_{X_1}^{-1}(u_1), \dots, F_{X_d}^{-1}(u_d)\right)$$

Stock data

Stocks have a nasty tendancy to crash together [3], which manifests as strong lower tail dependence in the copula. t-copulas can model strong dependence in the tails, wheras normal copulas cannot.

In this project, the log-stock returns of sets of 10 stocks from 4 sectors are fit with univariate ARMA-GARCH models, to remove autocorrelation. The copulas are fit to the stationary residuals of these distributions.

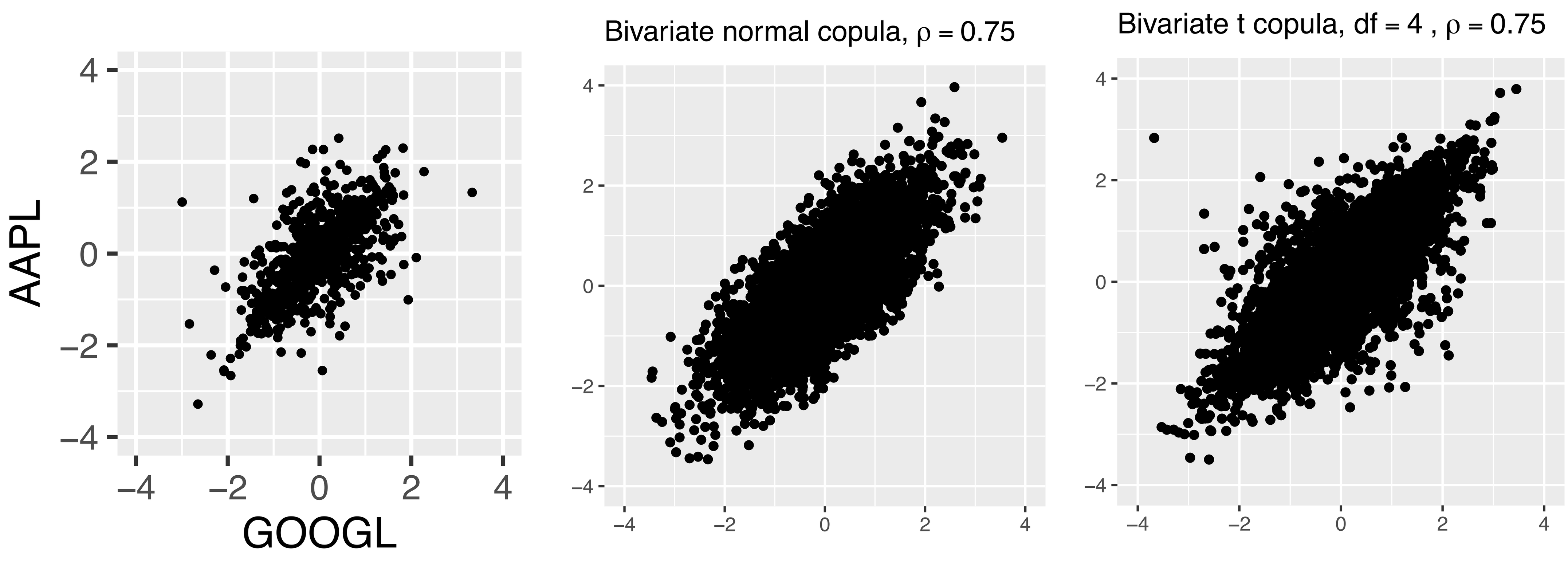

Figure 1: Normal scores of residuals of log-stock return from Apple and Google (left), and simulated bivariate normal (middle) and t_4 (right) copulas, $\rho = 0.75$.

Best model by BIC: Bifactor t-model

Designing tractable copulas remains a major challenge. This project uses factor normal and t-copula models, which reduces parameters by basing the covariance matrix of the copula on latent factors[4].

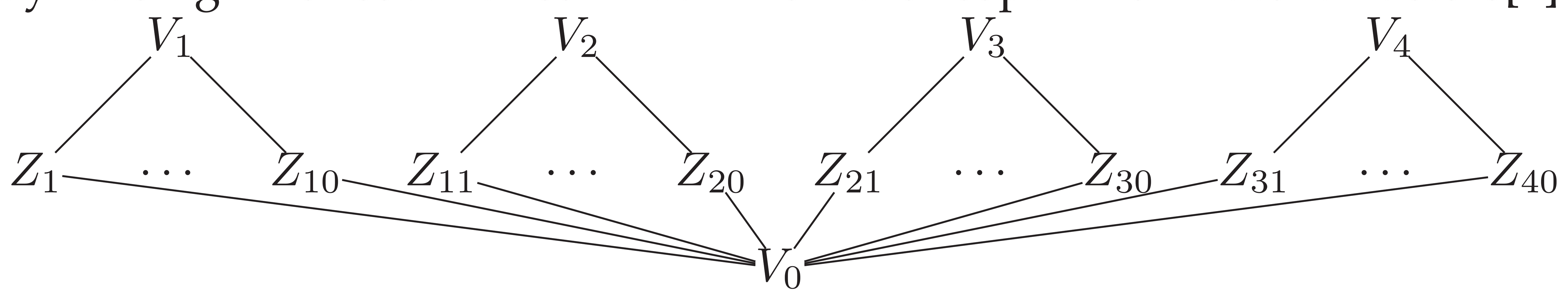

Figure 2: Structure of bifactor t-copula. $V_0, V_1, \dots V_4$ are latent random variables.

Here Z_i is the residual of the stock i transformed to $N(0,1)$-margins, where

$$Z_1 = a_1 V_0 + b_1 V_1 + \epsilon_1 V_{Z_1}, \qquad a_1^2 + b_1^2 + \epsilon_1^2 = 1$$

$$V_0, V_1, \dots, V_4, V_{Z_1}, \dots, V_{Z_k} \overset{iid}{\sim} N(0,1)$$

We fit the t-copula model to

$$\boldsymbol{S} = \begin{bmatrix} S_1 & \dots & S_{40} \end{bmatrix} = \frac{\begin{bmatrix} Z_1 & \dots & Z_{40} \end{bmatrix}}{\sqrt{U/\nu}}, \quad U \sim \chi^2_\nu$$

Then $\boldsymbol{S}$ is $\boldsymbol{t}_{\{1:d\}}(\nu, \boldsymbol{\Sigma})$-distributed where $\boldsymbol{\Sigma}_{ij} = \mathrm{COV}(Z_i, Z_j)$. Its DF is:

$$C(u_1, \dots, u_d) = \boldsymbol{T}_{\{1:d\}}\left(T^{-1}(u_1;\nu), \dots, T^{-1}(u_d;\nu);\, \nu, \boldsymbol{\Sigma}\right)$$

The t-copulas fit the data better in the tails than the normal models, but the fit is better in the upper tail than the lower tail. Next steps are to improve fit in the lower tail by using asymmetric structures.

References

[1] Abe Sklar. Fonctions de repartition a n dimensions et leurs marges. *Publications of the Institute of Statistics of the University of Paris*, 8:229–231, 1959.

[2] Berthold Schweizer, Edward F Wolff, et al. On nonparametric measures of dependence for random variables. *The annals of statistics*, 9(4):879–885, 1981.

[3] Fischer Black. The pricing of commodity contracts. *Journal of financial economics*, 3(1-2):167–179, 1976.

[4] Pavel Krupskii and Harry Joe. Factor copula models for multivariate data. *Journal of Multivariate Analysis*, 120:85–101, 2013.