Risk Management Project

Ashutosh
Harshit Goyal
Shaunak Gupta
Snigdha Jadhav

18th December 2023

1 Introduction

1.1 Objectives of this report

The recent Moody's downgrades and review announcements for 16 US banks have become a wake-up call to investors. The downgrades included several Trust banks, usually one of the lower-risk segments in the financial sector. The issue is higher interest rates: Trust banks have heavily depended on low-cost finance, and depositors are becoming much more active in moving cash balances. It highlights a broader issue – competition for funding is rising rapidly.

As a result, credit default rates are trending up, and likely to rise further over the next 12 months. S&P reported a provisional June 2023 default rate for US speculative grade ("High Yield") corporate bonds of 3.24% (up a hefty 74Bps from 2.5% in March), and this is currently projected to hit 4.25% by Q2 2024. They also project a pessimistic scenario where defaults hit 6.25% and an optimistic scenario where they drop back to 1.75%. The June increase is above the recent trend, so all the SP500 2024 projections – optimistic, median, and pessimistic – may be revised upwards in the second half of this year.

We have referred to [1] for this project motivation.
Over the years, many models under risk management have been developed to model these default risks. The project's main objective is to model and contrast a few of these industry-standard models with actual default risk probabilities that are adhered to by the sector.

This article may be understood as a model risk study in the context of latent variable models. Individual default probabilities and asset correlations are insufficient to determine the portfolio loss distribution since they do not fix the copula of the latent variables. For large portfolios of tens of thousands of counterparties, there remains considerable model risk. Risk managers who employ the latent variable methodology should be aware of this.
sectionTheoretical overview

1.2 Latent Variable Models

Referring to [2]
Consider a portfolio of m obligers and fix some time horizon T, typically one year. For 1 ¡= i ¡= m, let the random variable Y_i be the default indicator for obligor i at time T, taking values in 0; 1. We interpret the value 1 as default and 0 as non-default. At time t = 0 all obligors are assumed to be in a non-default state.

Let X = (X1,. . . . , Xm)' be an m-dimensional random vector with continuous marginal distributions representing the latent variables at time T and let $(D_1, ... , D_m)$ be a vector of deterministic cut-off levels. We call $(X_i, D_i)_{1<=i<=m}$ a latent variable model for the binary random vector Y = (Y1, . . . ,Ym)' if the following relationship holds:

$$Y_i = 1 \Leftrightarrow X_i \leq D_i \tag{1}$$

In the KMV model the latent variables Xi are assumed to be multivariate Gaussian and are interpreted as relative changes in the firm's asset value (so-called asset returns). For determining the thresholds Di an option pricing technique based on historical firm value data is used. The asset return correlations are calibrated by assuming that asset returns follow a factor model, where the underlying factors are interpreted as a set of macroeconomic variables.

1.3 Vasicek Model

Referring to [3]
The Vasicek (2002) model assumes that the asset value of a given obligor is given by the combined effect of a systematic and an idiosyncratic factor. It assumes an equi-correlated, Gaussian default structure. That is, each obligor i defaults if a certain random variable falls below a threshold, and these are all normal and equi-correlated. The asset value of the ith obligor at time t is therefore given by:

$$X_{it} = S_t\sqrt{\rho} + Z_{it}\sqrt{1-\rho} \tag{2}$$

Where S and Z are respectively the systematic and the idiosyncratic components and it can be proved that this is the asset correlation between two different obligors. The Vasicek model uses three inputs to calculate the probability of default (PD) of an asset class.

Vasicek applied to firms' asset values what had become the standard geometric Brownian motion model. Expressed as a stochastic differential equation,

$$dA_i = \mu_i A_i dt + \sigma_i A_i dx_i \tag{3}$$

Where is the value of the ith firm's assets, and are the drift rate and volatility of that value, and is a Wiener process or Brownian motion, i.e. a random walk in continuous time in which the change over any finite period is normally distributed with mean zero and variance equal to the length of the period, and changes in separate periods are independent of each other. Solving this stochastic differential equation one obtains the value of the ith firm's assets at time T as:

$$A_i = \exp\left(A(0) + \mu_i T - 0.5\sigma_i^2 T + \sigma_i\sqrt{T}X_i\right) \tag{4}$$

The ith firm defaults if $A_i(T) < B$, so the probability of such an event is

$$P[A_i(T) < B_i] = P[X_i < c_i] = N(c_i) = p* \tag{5}$$

where c_i is easily derived from the equation and N is the cumulative normal pdf. That is, the default of a single obligor happens if the value of a normal random variable happens to fall below a certain.

1.4 Merton Model

Referring to [4]
The Merton's model has been introduced in 1974.

1.4.1 Assumptions

1. The company is limited and the asset follows some stochastic process A_t at date_t.
2. Company's asset is financed by equity E_t and zero-coupon e_b to obligation D_t with face value D_T maturing at time $T > t$.

3. Market is assumed to be frictionless meaning that the value of the company's asset is equal to the sum of the value of debt obligation and the value of the equity at date t.

4. Company cannot pay dividend or issue new debt until the maturity T.

5. Default occurs if the company is not able to pay debt holders at maturity.

Under these assumptions, we can write down: $A_t = E_t + D_t$.
And, there are two possible scenarios at the maturity T.

1. The value of the company's assets exceeds that of liabilities, meaning that $A_T > D_T$. In this case, debt holders receive D_T, and shareholders $E_T = A_T - D_T$.

2. The value of the company's assets is less than the liabilities and the company is not able to meet its obligations. In other words, $A_T < D_T$. Thus, shareholders hand over control of the company to debt holders by exercising the limited liability option. Debt holders thus liquidate the company and distribute the revenues among them.

This can also be summarized by the two following equations.

$$E_T = \max(A_T - D_T, 0) \tag{6}$$

$$D_T = min(A_T, D_T) \tag{7}$$

1.4.2 Default Probability

The default occurs in Merton's model when at maturity T, the value of the company's asset A_T falls below the face value of the debt D_T. Then, the default probability (PD) in this way.

$$P_D = pro(A_T < D_T) \tag{8}$$

Merton assumes that the company's asset value A_t follows a geometric Brownian motion process, with risk-neutral dynamics given by the following stochastic differential equation.

$$\frac{dA_t}{A_t} = r\,dt + \sigma A_t\,dw_t. \tag{9}$$

where w_t is a standard Brownian motion under risk-neutral measure, r denotes the risk-free interest rate, and σ_A is the asset's return volatility. r captures the tendency of the asset's value while the volatility of the asset σ_A is assumed to be constant and captures fluctuations around the tendency.

1.5 Kealhofer, McQuown, and Vasicek (KMV) Model

Referring to [5]
The practical implementation of Merton's model has received considerable commercial attention in recent years. One of them is the KMV model which is a modified version of Merton's concept, varying from the original with a few aspects.

According to the preceding discussion, in Merton's model, a nominal value of a firm's obligation was considered as a terminal value for a firm's assets. "KMV Corporation has observed from a sample of several hundred companies that firms are generally more likely to default when their asset values reach a certain critical level somewhere between the value of total liabilities and the value of short-term debt. Therefore, in practice, using D alone as the threshold in the tail distribution might not be an accurate measure of the actual probability of default. KMV implements an additional step and refers to this critical threshold for defaulting as the Default Point." (Ong, 2005, p. 84). The ambiguity of the formal bankruptcy state and the situation, when assets value fall below the value of liabilities makes the theory of determining an accurate threshold level for default situation quite soft. For KMV model Default Point (DPT) is roughly approximated by the sum of all the Short Term Debt (STD) and half of the Long Term Debt (LTD)

$$DPT = STD + 0.5 * LTD \tag{10}$$

Another point is, that for practical reasons, before computing the probability of default, the KMV approach implements an intermediate phase of computation of an index called Distance to Default (DD) (Lu, 2008, p. 12). "It is defined as the distance between the expected assets value of the firm at the analysis horizon, (...) and the default point, normalized by the standard deviation of the future asset returns." (Ong, 2005, p. 84)

$$DD = \frac{(E(A_T) - DPT)}{\sigma} \tag{11}$$

Following that idea, to derive the probability of default for a particular firm, we must calculate the distance to default first. The probability of default for any time horizon is strongly related to DD. The larger DD, the smaller PD which means the less chance the company will default.

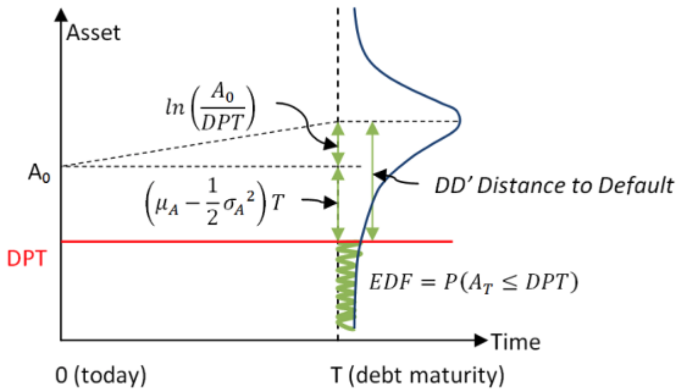

Figure 1: KMV Model

Calculating Distance to Default can be decomposed into two stages:

1. calculating absolute Distance to Default – DD'
2. calculating relative Distance to Default – DD.

Absolute Distance to Default (DD') is expressed (in percent of expected assets) as the distance between expected assets and Default Point (DPT). It can be displayed as a sum of the initial distance and the growth of that distance within the period T.

$$DD' = ln\frac{A_0}{DPT} + (\mu_A - 0.5\sigma_A^2)T \tag{12}$$

As in the pure Merton concept, in the KMV model, μ_A is no longer a risk-free rate but the expected rate of the return of the firm's asset and DPT is the Default Point instead of the nominal value D (the face value of the debt). While the rate of return is normally distributed, consequently the future value of investment (or effective yield of return) is distributed log normally. The relation between those two distributions is explained with (24), where, as previously, μ_A is the drift rate (expected rate of return) and σ_A is the volatility of the underlying (The Professional..., 2004).

$$ln(\frac{S_T}{S})\ N((\mu_A - 0.5\sigma_A^2)T, \sigma_A\sqrt{T}) \tag{13}$$

Dividing absolute value DD' with calibrated (according to T – usually annualized) volatility of assets, we can calculate DD in relative terms as a multiplier of standard deviation

$$DD = d_2 = \frac{ln\frac{A_0}{DPT} + (\mu_A - 0.5\sigma_A^2)T}{\sigma_A\sqrt{T}} \tag{14}$$

It is easy to notice, that such estimation of Distance to Default is very similar to d2 (considering mentioned above replacement of r with μ_A and D with DPT). "The similarity is not an accident and is the result of a relationship between the risk – neutral probability and the actual probability. The actual probability uses the expected return of the assets in the drift term, while the risk-neutral probability uses the risk free rate r." (Ong, 2005, p. 86).

Because of well-known problem of fatter tails in real credit loss distribution, that type of estimation (even with lognormal distribution instead of normal) is underappreciated. In that situation, one more distinguishing feature of KMV model is, that it operates on the historical set of frequencies of default rather than on theoretical normal or log-normal distribution. Consequently, in KMV model Probability of Default (PD) is replaced with Expected Default Frequency (EDF). "Using historical information about a large sample of firms, including firms that have defaulted, one can track, for each time horizon, the proportion of firms of a given ranking (...) that defaulted after one year." (Crouchy, Galai, Mark, 2006, p. 277). An example of that dependence is shown at Figure 9. For DD = 3, the Expected Default Frequency is equal to 40 basic points. That means, that according to database analysis, 0.4% of registered firms with DD = 3 defaulted after one year. At the same time, for DD = 1, EDF grows to 120 b.p. = 1.2%.

The simplicity of using EDF vs. DD curve makes the concept of KMV model very comprehensive and easy to implement.

1.6 Copulas

Referring to [2]
Copulas are simply the joint distribution functions of random vectors with standard uniform marginal distributions. Their value in statistics is that they provide a way of understanding how marginal distributions of single risks are coupled together to form joint distributions of groups of risks; that is, they provide a way of understanding the idea of statistical dependence.

There are two principal ways of using the copula idea. We can extract copulas from well-known multivariate distribution functions. We can also create new multivariate distribution functions by joining arbitrary marginal distributions together with copulas. These ideas are

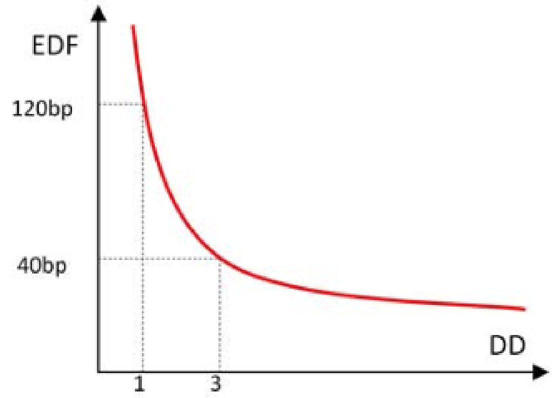

Figure 2: Correlation between Distance to Default and Expected Default Frequency

summarised in the following proposition, known as Sklar's Theorem; see Nelsen (1999) for proof.

To understand that the use of the Gaussian copula leads to structurally equivalent models we introduce a formal definition of equivalence for latent variable models and present a simple new result. Let $(X_i, D_i)_{1<=i<=m}$ and $(\tilde{X}_i, \tilde{D}_i)_{1<=i<=m}$ be two latent variable models generating default indicator vectors Y and $\tilde{Y}$. The models are called equivalent if $Y \stackrel{d}{=} \tilde{Y}$.

Thus two models are equivalent if they give rise to the same default indicator distribution, which means of course that the distribution of the number of defaults in the portfolio will be the same.

A sufficient condition for two latent variable models to be equivalent is that individual default probabilities are the same in both models and the copulas of the latent variables are the same. Formally we have the following, which is proved in Frey and McNeil (2001).

There are many alternative copulas to the Gaussian. We choose to work with the copulas that are implicit in the kinds of multivariate distributions that might be considered natural alternative models for asset values and asset returns. It would also be possible to work with families of simple closed-form parametric copulas such as the Archimedean family(Nelsen 1999).

A popular family of distributions for modeling financial market returns is the family of multivariate normal mixture models. When relaxing the assumption of multivariate normality for asset returns it seems natural to look at this family, which contains such distributions as the multivariate t and the hyperbolic.

2 Method Used

We have decided to make use of

1. Combined Merton and Vasicek Model

2. KMV model

3. Clayton and Gumbel Copula

to compare the probability of default among themselves and to the closeness of default probability to that in the industry.

Using first N values for mean return and volatility calculations, we have employed the rolling and recursive window technique to obtain default distance and subsequently probabilities. These have been updated in computations through the use of the aforementioned looping technique. We have used this technique using [6]

2.1 Combined Merton and Vasicek Mode

Here we exploit the use of Merton and Vasicek model to get default probabilities. We have given equal weights to either of the models i.e., DP = 0.5*(DP from Merton model) + 0.5*(DP from the Vasicek model). Even though Vasicek model is used for interest rate modelling we can use this model for default rate and probabilities calculation.

Merton Probability

$$d1 = \frac{\log\left(\frac{\text{total assets}}{\text{debt value}}\right) + 0.5 \times \text{Standard Deviation of Market Value} \times \text{time\_to\_maturity}}{\text{Standard Deviation of Market Value} \times \sqrt{\text{time\_to\_maturity}}}$$

$$d2 = d1 - \text{Standard Deviation of Market Value} \times \sqrt{\text{time\_to\_maturity}}$$

$$\text{Merton probability} = \Phi(-d2)$$

where Φ is the cumulative distribution function of the standard normal distribution.

Vasicek Probability

$$\text{Vasicek probability} = \frac{\log\left(\frac{\text{debt value}}{\text{equity value}}\right) + \left(\text{Volatility of equity}^2 \times \text{time to maturity}\right)}{\text{Volatility of equity} \times \sqrt{\text{time to maturity}}}$$

Finally we use weighted average.

$$\text{Combined probability} = 0.5 \times \text{Merton probability} + 0.5 \times \text{Vasicek probability}$$

We have then set a probability of 0.7 as a threshold and set alerts that would pop up when the default probability of a stock goes above the threshold in a given quarter.

2.2 KMV Model

Here, we exploit the use of the KMV model and its valuation for default probabilities. We have used the same logic as given in the above KMV section.

Default distance using KMV

$$d = \frac{\log\left(\frac{\text{Market Value}}{0.5\times\text{Long Term Debt}+\text{Short Term Debt}}\right) + \left(\text{mean return} - \frac{\text{return volatility}}{2}\right) \times \text{time\_to\_maturity}}{\text{return volatility} \times \sqrt{\text{time to maturity}}}$$

$$Probability = \min\left(\frac{d}{40}, 1\right)$$

We have used a simple probability logic where every default distance will lead to a probability of 1 for default. This is as per the KMV section in the report.

2.3 Copulas

We have used a similar concept for Copula based models.

2.3.1 Gumbel Copula

The Gumbel copula is a copula that allows any specific level of (upper) tail dependency between individual variables. It is an Archimedean copula, and exchangeable.

The Gumbel copula formula is given by:

$$C(u, v; \theta) = \exp\left(-\left((-\log u)^\theta + (-\log v)^\theta\right)^{\frac{1}{\theta}}\right)$$

where:

$$0 < \theta \leq 1$$
$$0 < u, v < 1$$

2.3.2 Calyton Copula

The Clayton copula is a copula that allows any specific non-zero level of (lower) tail dependency between individual variables. It is an Archimedean copula, and exchangeable.

The Clayton copula formula is given by:

$$C(u, v; \theta) = \left(u^{-\theta} + v^{-\theta} - 1\right)^{-1/\theta}$$

where u and v are the marginal distribution functions, and θ is the copula parameter.

Probability of Default:
Using functions of copula we fit the model in a rolling fashion and get transformed sample using the model fit. We then define the default.

P = $transformed\_samples$('Market Value') * $transformed\_samples$('Total Debt')

3 Results

Given that we have used stock constituents of SP500, we are expecting very low or no defaults as SP500 is made up of the biggest 500 companies. We are also a bit skeptical on the accuracy of mixed model and KMV model as these models have assumptions which do not meet the real markers and it's imperfections. We are hoping to see that copulas will give us good results for default probabilities taking into account the multivariate nature . Therefore we have set a lower threshold of 0.6 for copulas.

We compared the results of each model with the industrial default risk of SP500. According to [7] the default risk of SP500 ranges from 3-7% as per historical and future references via SP500 officially. This range will serve as a industrial benchmark.

3.1 Combined Merton and Vasicek Mode

In view of the assumptions in both of these models, it is expected that the probabilities from these models will be very off from the industrial range.

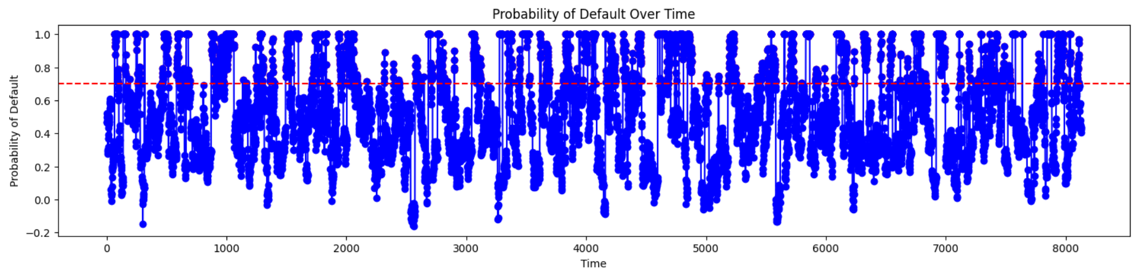

Figure 3: Probability of Default over time for 226 stock using mixed model

```
Alert: Probability of default for AMD exceeds 0.7. Quarter: 2015Q2
Alert: Probability of default for AMD exceeds 0.7. Quarter: 2015Q3
Alert: Probability of default for AMD exceeds 0.7. Quarter: 2015Q4
Alert: Probability of default for AMD exceeds 0.7. Quarter: 2016Q1
Alert: Probability of default for ALK exceeds 0.7. Quarter: 2020Q1
Alert: Probability of default for ALK exceeds 0.7. Quarter: 2020Q2
Alert: Probability of default for ALK exceeds 0.7. Quarter: 2020Q3
Alert: Probability of default for APA exceeds 0.7. Quarter: 2020Q2
Alert: Probability of default for APA exceeds 0.7. Quarter: 2020Q4
Alert: Probability of default for APA exceeds 0.7. Quarter: 2021Q1
Alert: Probability of default for APA exceeds 0.7. Quarter: 2021Q2
Alert: Probability of default for APA exceeds 0.7. Quarter: 2021Q3
Alert: Probability of default for CAG exceeds 0.7. Quarter: 2019Q1
Alert: Probability of default for TAP exceeds 0.7. Quarter: 2018Q2
Alert: Probability of default for TAP exceeds 0.7. Quarter: 2018Q3
Alert: Probability of default for TAP exceeds 0.7. Quarter: 2018Q4
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2020Q1
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2020Q4
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2021Q1
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2021Q2
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2021Q3
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2021Q4
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2022Q1
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2022Q2
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2022Q3
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2022Q4
Alert: Probability of default for DAL exceeds 0.7. Quarter: 2023Q1
Alert: Probability of default for FDX exceeds 0.7. Quarter: 2020Q2
Alert: Probability of default for HAL exceeds 0.7. Quarter: 2018Q2
Alert: Probability of default for HAL exceeds 0.7. Quarter: 2018Q3
Alert: Probability of default for HAL exceeds 0.7. Quarter: 2018Q4
Alert: Probability of default for IP exceeds 0.7. Quarter: 2015Q3
Alert: Probability of default for IP exceeds 0.7. Quarter: 2015Q4
Alert: Probability of default for IP exceeds 0.7. Quarter: 2018Q4
Alert: Probability of default for IP exceeds 0.7. Quarter: 2019Q2
Alert: Probability of default for IP exceeds 0.7. Quarter: 2019Q3
Alert: Probability of default for IP exceeds 0.7. Quarter: 2020Q1
Alert: Probability of default for IP exceeds 0.7. Quarter: 2020Q3
Alert: Probability of default for KR exceeds 0.7. Quarter: 2018Q3
Alert: Probability of default for KR exceeds 0.7. Quarter: 2018Q4
Alert: Probability of default for KR exceeds 0.7. Quarter: 2019Q1
```

Figure 4: Default results from mixed model

Comments: The average probability of default was around 52.33%. This was expected to be way off in view of the assumptions made for these models to be theoretically easy.

3.2 KMV model

Similar to the mixed model. KMV proposes a lot of assumptions for modeling which sets it apart from the industrial range. The expectation here is similar to that of the mixed model where the probability of default from KMV will not be aligned with the industrial standards.

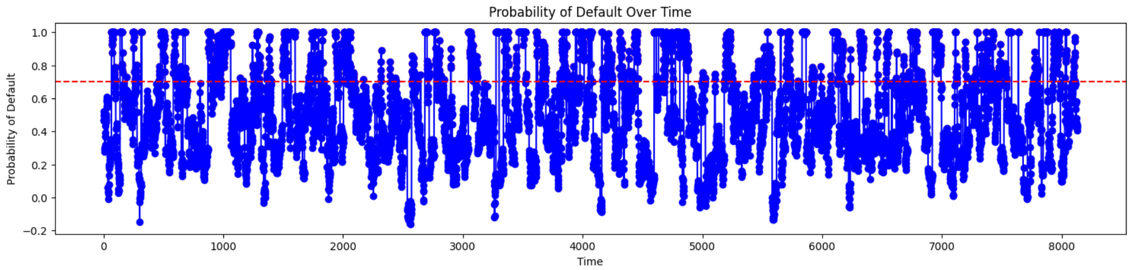

Figure 5: Probability of Default over time for 226 stock using KMV model

```
Alert: Probability of default for APD exceeds 0.7. Quarter: 2014Q3
Alert: Probability of default for APD exceeds 0.7. Quarter: 2014Q4
Alert: Probability of default for APD exceeds 0.7. Quarter: 2015Q1
Alert: Probability of default for APD exceeds 0.7. Quarter: 2015Q2
Alert: Probability of default for APD exceeds 0.7. Quarter: 2015Q3
Alert: Probability of default for APD exceeds 0.7. Quarter: 2015Q4
Alert: Probability of default for APD exceeds 0.7. Quarter: 2016Q1
Alert: Probability of default for APD exceeds 0.7. Quarter: 2016Q2
Alert: Probability of default for APD exceeds 0.7. Quarter: 2016Q3
Alert: Probability of default for APD exceeds 0.7. Quarter: 2016Q4
Alert: Probability of default for APD exceeds 0.7. Quarter: 2017Q1
Alert: Probability of default for APD exceeds 0.7. Quarter: 2017Q2
Alert: Probability of default for APD exceeds 0.7. Quarter: 2017Q3
Alert: Probability of default for APD exceeds 0.7. Quarter: 2017Q4
Alert: Probability of default for APD exceeds 0.7. Quarter: 2018Q1
Alert: Probability of default for APD exceeds 0.7. Quarter: 2018Q2
Alert: Probability of default for APD exceeds 0.7. Quarter: 2018Q3
Alert: Probability of default for APD exceeds 0.7. Quarter: 2018Q4
Alert: Probability of default for APD exceeds 0.7. Quarter: 2019Q1
Alert: Probability of default for APD exceeds 0.7. Quarter: 2019Q2
Alert: Probability of default for APD exceeds 0.7. Quarter: 2019Q3
Alert: Probability of default for APD exceeds 0.7. Quarter: 2019Q4
Alert: Probability of default for APD exceeds 0.7. Quarter: 2020Q1
Alert: Probability of default for APD exceeds 0.7. Quarter: 2020Q2
Alert: Probability of default for APD exceeds 0.7. Quarter: 2020Q3
Alert: Probability of default for APD exceeds 0.7. Quarter: 2020Q4
Alert: Probability of default for APD exceeds 0.7. Quarter: 2021Q1
Alert: Probability of default for APD exceeds 0.7. Quarter: 2021Q2
Alert: Probability of default for APD exceeds 0.7. Quarter: 2021Q3
Alert: Probability of default for APD exceeds 0.7. Quarter: 2021Q4
Alert: Probability of default for HON exceeds 0.7. Quarter: 2014Q3
Alert: Probability of default for HON exceeds 0.7. Quarter: 2014Q4
Alert: Probability of default for HON exceeds 0.7. Quarter: 2015Q1
Alert: Probability of default for HON exceeds 0.7. Quarter: 2015Q2
Alert: Probability of default for HON exceeds 0.7. Quarter: 2015Q3
Alert: Probability of default for HON exceeds 0.7. Quarter: 2015Q4
Alert: Probability of default for HON exceeds 0.7. Quarter: 2016Q1
Alert: Probability of default for HON exceeds 0.7. Quarter: 2016Q2
Alert: Probability of default for HON exceeds 0.7. Quarter: 2016Q3
Alert: Probability of default for HON exceeds 0.7. Quarter: 2016Q4
Alert: Probability of default for HON exceeds 0.7. Quarter: 2017Q1
```

Figure 6: Default results from KMV

Comments: The average probability of default was around 53% similar to that of the mixed model . This also aligns with a presumption that the default probability to be way off in view of the assumptions made for KMV model to be theoretically easy.

3.3 Gumbel Copula

Our presumption is to see good results from Gumbel copulas which takes into account copula samples.

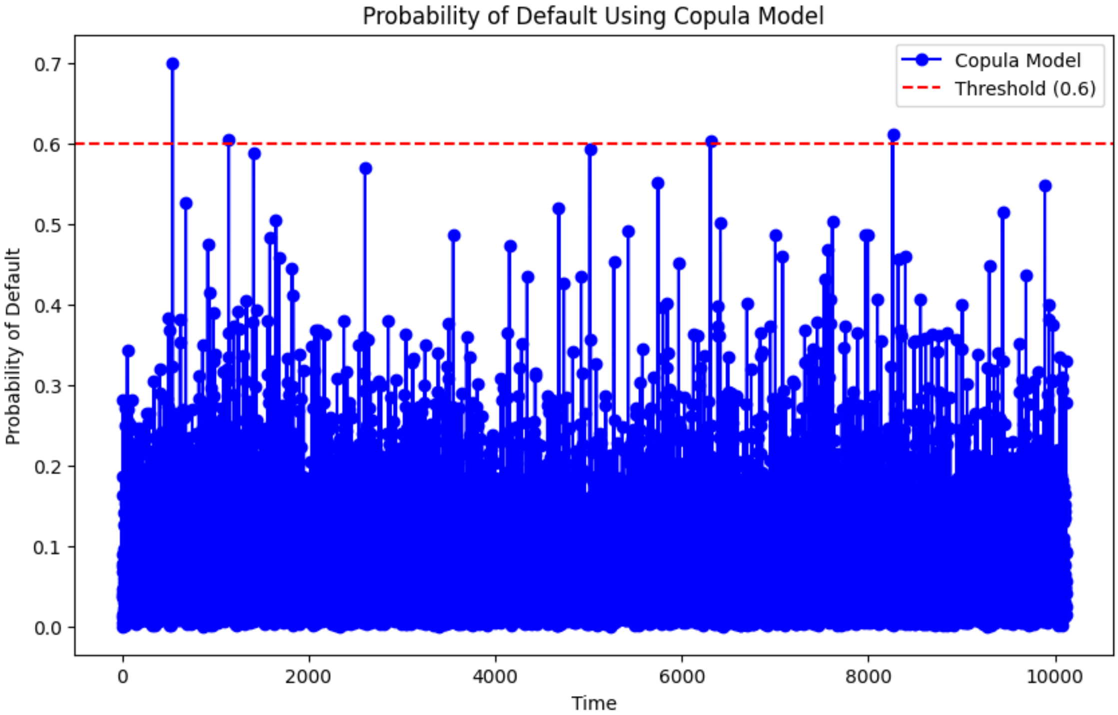

Figure 7: Probability of Default over time for 226 stock using Gumbel model

```
Alert: Probability of default for ADM exceeds 0.6. Quarter: 2022Q1
Alert: Probability of default for CTAS exceeds 0.6. Quarter: 2012Q1
Alert: Probability of default for FCX exceeds 0.6. Quarter: 2019Q2
Alert: Probability of default for DLTR exceeds 0.6. Quarter: 2017Q4
```

Figure 8: Default results from Gumbel

Comment: Gumbel model gave 4 alerts which has been a really good improvement from mixed and KMV model. The probability of default from Gumbel Copula was around 8.3% which is not perfect but a really good estimation of SP500 default risk.

3.4 Clayton Copula

Our presumption is to see good results from Clayton copulas as well. Similar to Gumbel copula model.

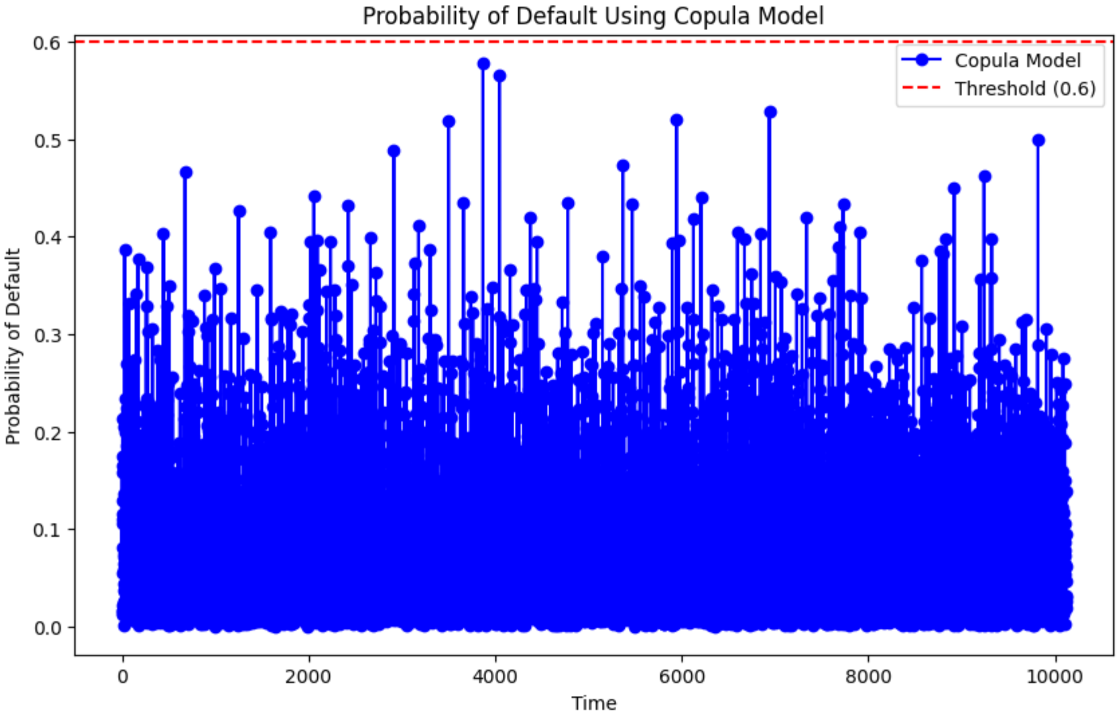

Figure 9: Probability of Default over time for 226 stock using Gumbel model

Comment: Clayton model gave 0 alerts which was really impressive and something that we were expecting from SP500. The probability of default from Clayton copula was around 8.2% which again is a really good estimation of industrial standard risk.

References

[1] Default rate forecast 2023/24: Us speculative grade borrowers and us leveraged loans. [Online]. Available: https://www.creditbenchmark.com/default-rate-forecast-2023-24-us-speculative-grade-borrowers-and-us-leveraged-loans/

[2] M. A. N. R¨udiger Frey, Alexander J. McNeil, "Copulas and credit models," *Academic Papers*, 2001. [Online]. Available: https://statmath.wu.ac.at/~frey/publications/copula-credit.pdf

[3] Bank of england: Centre for central banking studies. [Online]. Available: https://www.bankofengland.co.uk/-/media/boe/files/ccbs/resources/modelling-credit-risk

[4] Delft university of technology: Lmv model. [Online]. Available: https://courses.edx.org/c4x/DelftX/TW3421x/asset/Week5_PD2_3.pdf

[5] A. W. John Hull, Izzy Nelken, "Merton's model, credit risk, and volatility skews," *Academic Papers*, 2004. [Online]. Available: http://www-2.rotman.utoronto.ca/~hull/downloadablepublications/mertonsmodelandvolatilityskews.pdf

[6] D. X. Shihao Gu, Bryan Kelly, *Empirical Asset Pricing via Machine Learning.* Academic Press, 2018. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3159577

[7] Default risk broadly ticks up for sp 500 companies at end of june. [Online]. Available: https://www.spglobal.com/marketintelligence/en/news-insights/latest-news-headlines/default-risk-broadly-ticks-up-for-s-p-500-companies-at-end-of-june-76344416#:~:text=The%20median%20probability%20of%20default,Market%20Intelligence's%20Risk%20Gauge%20model.

Python code

```
# -*- coding: utf-8 -*-
"""Risk_Management_Project.ipynb

Automatically generated by Colaboratory.

Original file is located at
    https://colab.research.google.com/drive/1hQt5dHsmjRwXIFz0Xul_Xz30bYq43sF9
"""

import warnings
warnings.filterwarnings('ignore')

"""1. Read Data"""

# read csv file
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
data = pd.read_csv('Data_Project.csv')
data = data.set_index(['Ticker Symbol', 'Calendar Data Year and Quarter'], drop=True, verify_integrity=True)
data = data.groupby(level=0).filter(lambda x: len(x) >= 35)
data.groupby(level=0).size()
data.drop(columns=['Active/Inactive Status Marker'], inplace=True)
data = data.rename(columns={'Long-Term Debt Due in One Year': 'Short_Term_Debt'})
data = data.rename(columns={'Long-Term Debt - Total': 'Long_Term_Debt'})
data.info(verbose=True, show_counts=True)

"""2. Merton and Vasicek Combined"""

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

rolling_window_size = 10

new_data = data.copy()

new_data['Total_Debt'] = new_data['Short_Term_Debt'] + new_data['Long_Term_Debt']
new_data['Asset_Std'] = np.nan
new_data['KMV_Distance'] = np.nan
new_data['KMV_Probability_of_Default'] = np.nan

time_to_maturity = 1
alerts = []

probability_of_default_list = []

for ticker in new_data.index.get_level_values('Ticker Symbol').unique():
    for i in range(len(new_data.loc[ticker]) - rolling_window_size):
        subset = new_data.loc[ticker].iloc[:i+rolling_window_size]

```

```python

        equity_value = subset.iloc[-1]['Market Value - Total'] - 0.5 * subset.iloc[-1]['Long_Term_Debt'] - subset.iloc[-1]['Short_Term_Debt']
        total_assets = subset.iloc[-1]['Market Value - Total']
        debt_value = subset.iloc[-1]['Total_Debt']

        d1 = (np.log(total_assets / debt_value) + (0.5 * subset['Market Value - Total'].pct_change().std() ** 2) * time_to_maturity) / (subset['Market Value - Total'].pct_change().std() * np.sqrt(time_to_maturity))
        d2 = d1 - subset['Market Value - Total'].pct_change().std() * np.sqrt(time_to_maturity)

        merton_prob = norm.cdf(-d2)

        equity_volatility = subset['Market Value - Total'].pct_change().std()
        vasicek_prob = norm.cdf((np.log(debt_value / equity_value) + (equity_volatility ** 2) * time_to_maturity) / (equity_volatility * np.sqrt(time_to_maturity)))

        combined_prob = 0.5 * merton_prob + 0.5 * vasicek_prob

        new_data.at[(ticker, subset.index[-1]), 'KMV_Probability_of_Default'] = combined_prob

        probability_of_default_list.append(combined_prob)

        if combined_prob > 0.7:
            quarter = new_data.index.get_level_values('Calendar Data Year and Quarter')[i + rolling_window_size - 1]
            print(f"Alert: Probability of default for {ticker} exceeds 0.7. Quarter: {quarter}")
            alerts.append(f"Alert: Probability of default for {ticker} exceeds 0.7. Quarter: {quarter}")

import numpy as np
import matplotlib.pyplot as plt

plt.figure(figsize=(15, 10))

plt.subplot(3, 1, 1)
plt.plot(probability_of_default_list, label='Probability of Default', marker='o', color='b')

above_threshold_indices = np.where(np.array(probability_of_default_list) > 0.7)[0]
plt.scatter(above_threshold_indices, np.array(probability_of_default_list)[above_threshold_indices], color='r')

plt.axhline(y=0.7, color='r', linestyle='--', label='Threshold (0.7)')
plt.title('Probability of Default Over Time')
plt.xlabel('Time')
plt.ylabel('Probability of Default')
plt.legend().set_visible(False)

plt.tight_layout()
plt.show()
print(100*np.mean(probability_of_default_list))

```

```python
"""3. KMV model"""

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

rolling_window_size = 10

new_data = data.copy()

new_data['Total_Debt'] = new_data['Short_Term_Debt'] + new_data['Long_Term_Debt']
new_data['Asset_Std'] = np.nan
new_data['KMV_Distance'] = np.nan
new_data['KMV_Probability_of_Default'] = np.nan

time_to_maturity = 1
alerts = []

probability_of_default_list = []

for ticker in new_data.index.get_level_values('Ticker Symbol').unique():
    for i in range(len(new_data.loc[ticker]) - rolling_window_size):
        subset = new_data.loc[ticker].iloc[:i+rolling_window_size]
        std = subset['Market Value - Total'].pct_change().std()
        mean = subset['Market Value - Total'].pct_change().mean()

        d = (np.log(subset.iloc[-1]['Market Value - Total'] / (0.5 * subset.iloc[-1]['Long_Term_Debt'] + subset.iloc[-1]['Short_Term_Debt'])) + (mean - std/2) * time_to_maturity) / (std * np.sqrt(time_to_maturity))
        p = min(d / 40, 1)

        new_data.at[(ticker, subset.index[-1]), 'KMV_Probability_of_Default'] = p

        probability_of_default_list.append(p)

        if p > 0.7:
            quarter = new_data.index.get_level_values('Calendar Data Year and Quarter')[i + rolling_window_size - 1]
            print(f"Alert: Probability of default for {ticker} exceeds 0.7. Quarter: {quarter}")
            alerts.append(f"Alert: Probability of default for {ticker} exceeds 0.7. Quarter: {quarter}")

import numpy as np
import matplotlib.pyplot as plt

plt.figure(figsize=(15, 10))
plt.subplot(3, 1, 1)
plt.plot(probability_of_default_list, label='Probability of Default', marker='o', color='b')

above_threshold_indices = np.where(np.array(probability_of_default_list) > 0.7)[0]
plt.scatter(above_threshold_indices, np.array(probability_of_default_list)[above_threshold_indices], color='r')

```

```python
plt.axhline(y=0.7, color='r', linestyle='--', label='Threshold (0.7)')
plt.title('Probability of Default Over Time')
plt.xlabel('Time')
plt.ylabel('Probability of Default')
plt.legend().set_visible(False)

plt.tight_layout()
plt.show()
print(100*np.mean(probability_of_default_list))

"""4. Copula Model"""

debt_changes = new_data['Total_Debt'].pct_change().dropna()
value_changes = new_data['Market Value - Total'].pct_change().dropna()
changes_data = pd.DataFrame({'Total_Debt_Change': debt_changes, 'Value_Change': value_changes})
merged_data = new_data.merge(changes_data, left_index=True, right_index=True)
merged_data.dropna(subset=['Total_Debt_Change', 'Value_Change'], inplace=True)

copula_family = 'gumbel'
market_value_changes = merged_data['Market Value - Total'].pct_change().dropna()
total_debt_changes = merged_data['Total_Debt'].pct_change().dropna()

market_value_changes = np.clip(market_value_changes, 0, 1)
total_debt_changes = np.clip(total_debt_changes, 0, 1)

copula = Gumbel()
copula.fit(np.column_stack((market_value_changes.values, total_debt_changes.values)))

copula_samples = copula.sample(len(merged_data))

asset_cdfs = {'Market Value - Total': Beta(a=2, b=5), 'Total_Debt': Beta(a=2, b=5)}
transformed_samples = pd.DataFrame({
    asset: asset_cdfs[asset].ppf(copula_samples[:, i])
    for i, asset in enumerate(['Market Value - Total', 'Total_Debt'])
})

threshold = 0.6

probability_of_default_copula = []
for i in range(len(merged_data)):
    p_default = transformed_samples.iloc[i]['Market Value - Total'] * transformed_samples.iloc[i]['Total_Debt']
    probability_of_default_copula.append(p_default)

    if p_default > threshold:
        ticker = merged_data.index.get_level_values('Ticker Symbol')[i]
        quarter = merged_data.index.get_level_values('Calendar Data Year and Quarter')[i]
        print(f"Alert: Probability of default for {ticker} exceeds {threshold}. Quarter: {quarter}")

plt.figure(figsize=(10, 6))
plt.plot(probability_of_default_copula, label='Copula Model', marker='o',
```

```
    color='b')
plt.axhline(y=threshold, color='r', linestyle='--', label=f'Threshold ({threshold})')
plt.title('Probability of Default Using Copula Model')
plt.xlabel('Time')
plt.ylabel('Probability of Default')
plt.legend()
plt.show()
print(100*np.mean(probability_of_default_copula))

from copulas.bivariate import *
from copulas import *
from scipy.stats import beta as Beta
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

copula_family = 'clayton'
market_value_changes = merged_data['Market Value - Total'].pct_change().dropna()
total_debt_changes = merged_data['Total_Debt'].pct_change().dropna()

market_value_changes = np.clip(market_value_changes, 0, 1)
total_debt_changes = np.clip(total_debt_changes, 0, 1)

copula = Clayton()
copula.fit(np.column_stack((market_value_changes.values, total_debt_changes.values)))

copula_samples = copula.sample(len(merged_data))

asset_cdfs = {'Market Value - Total': Beta(a=2, b=5), 'Total_Debt': Beta(a=2, b=5)}
transformed_samples = pd.DataFrame({
    asset: asset_cdfs[asset].ppf(copula_samples[:, i])
    for i, asset in enumerate(['Market Value - Total', 'Total_Debt'])
})

threshold = 0.6

probability_of_default_copula = []
for i in range(len(merged_data)):
    p_default = transformed_samples.iloc[i]['Market Value - Total'] * transformed_samples.iloc[i]['Total_Debt']
    probability_of_default_copula.append(p_default)

    if p_default > threshold:
        ticker = merged_data.index.get_level_values('Ticker Symbol')[i]
        quarter = merged_data.index.get_level_values('Calendar Data Year and Quarter')[i]
        print(f"Alert: Probability of default for {ticker} exceeds {threshold}. Quarter: {quarter}")

plt.figure(figsize=(10, 6))
plt.plot(probability_of_default_copula, label='Copula Model', marker='o', color='b')
plt.axhline(y=threshold, color='r', linestyle='--', label=f'Threshold ({threshold})')
plt.title('Probability of Default Using Copula Model')
```

```python
plt.xlabel('Time')
plt.ylabel('Probability of Default')
plt.legend()
plt.show()
print(100*np.mean(probability_of_default_copula))
```