DEFAULT PROBABILITY PREDICTION WITH STATIC MERTON-D-VINE COPULA MODEL

Václav Klepáč[1]

[1] *Mendel University in Brno*

ABSTRACT

We apply standard Merton and enhanced Merton-D-Vine copula model for the measurement of credit risk on the basis of accounting and stock market data for 4 companies from Prague Stock Exchange, in the midterm horizon of 4 years. Basic Merton structural credit model is based on assumption that firm equity is European option on company assets. Consequently enhanced Merton model take in account market data, dependence between daily returns and its volatility and helps to evaluate and project the credit quality of selected companies, i.e. correlation between assets trajectories through copulas. From our and previous results it is obvious that basic Merton model significantly underestimates actual level, i.e. offers low probabilities of default. Enhanced model support us with higher simulated probability rates which mean that capturing of market risk and transferring it to credit risk estimates is probably a good way or basic step in enhancing Merton methodology.

KEY WORDS

Merton model, default risk, d-vine copula, probability, ARMA-GARCH

JEL CODES

C15, C53

1 INTRODUCTION

For the purpose of the quantification of the probability of the debtor's settlement of payment obligations in course of time, the following probabilities can be quantified: the particular default probabilities, the defaults for debt instruments in portfolios and defaults depending on the default of another subject. Basic models, which were created for the purpose of measurement of the risk of bankruptcy and financial health of companies, include the models based purely on accounting data and statistical methods.

One of the first authors who used basic statistical techniques in financial distress, e.g. bankruptcy prediction were Beaver (1966) with univariate analysis and Altman (1968) who used Multiple discriminant analysis (MDA), in that he computed an individual firm's discriminant score using a set of financial and economic ratios. Probably due to the huge demand coming from the financial sector in the beginning of 1980s more advanced estimation methods, such as Ohlson's logit (1980) and Zmijewski's probit (1984), were employed. Compared to the MDA the logit model was easier to understand since the logistic score, taking a value between 0 and 1, was interpretable in a probabilistic way.

Credit authorities need to estimate the probability of return of money lent (Credit risk). According to Míšek (2006), methods for quantification of so-called defaults (inability to repay obligations, bankruptcy) were formed. Another branches to statistical or data mining models are reduced type models (based on market data: bond and credit default swap prices etc.) and structural risk models (Merton model, 1973; Longstaff and Schwartz, 1995 etc.). Without any doubt these models are included among the influential methods for the credit risks measurement, which is used even in rating agencies (like KMV Moody's methodology).

Based on results from Klepáč (2014) and above stated authors we have realized that the structural models use the approaches based on option assessment. In these models, the value of assets (of the company) – after having exceeded given level – will cause the default of the company. The default probabilities can be measured on the basis of the distance between the market value of the company and the level of maximally financially manageable debts (non-leading to defaults). The knowledge of the probability of default enables in further applications estimation of credit spread (i.e. surcharge on risk-free interest rate) which will compensate the creditor's possible financial loss connected with the run default risk.

Therefore, the crucially important characteristic of credit models is the barrier which determines the default limit. Any change of expectations of a company's future means that especially the shares will react intensively to these changes. It is given by the fact that shareholders are the last subjects who claims will be discharged out of the company's remaining value in case of the company's default. That is why the latest information on company's fundaments are theoretically reflected both in the shares traded on the stock exchange and in the asset value which is generated via the above mentioned shares. Therefore, the market asset value includes both the future prospects of the company and the relevant information on the sector and economy, in which the company operates. Volatility in the sense of time-varying standard deviation of market price reflects the company's trade risk and the relevant sector risk.

Contribution of the article lies in comparison of the different ways of probability of default estimation. That is proposed by standard Merton model. We propose the novel methodological generalization of Merton model by high dimensional copulae for capturing market risk and transferring its character to default risk estimates with correlation of the traded equity in mind. Thus we partly continue, thereby with enhanced models, on the basis of results presented in Klepáč (2014).

2 THEORETICAL BACKGROUND: MERTON MODEL

Financial theory and risk management uses many Lévy procesess for modelling of asset returns and either for credit risk measurement. Probably the most famous is geometric Brownian motion used in Merton model as driving process of assets, which holds this form as stated in Schoutens and Cariboni (2009) or in Klepáč (2014):

$$dV_t = \mu V_t dt + \sigma V_t dW_t, \quad V_0 > 0, \qquad (1)$$

where $W = \{W_t, t > 0\}$ is standard Brownian motion, μ is the so-called drift parameter (mean

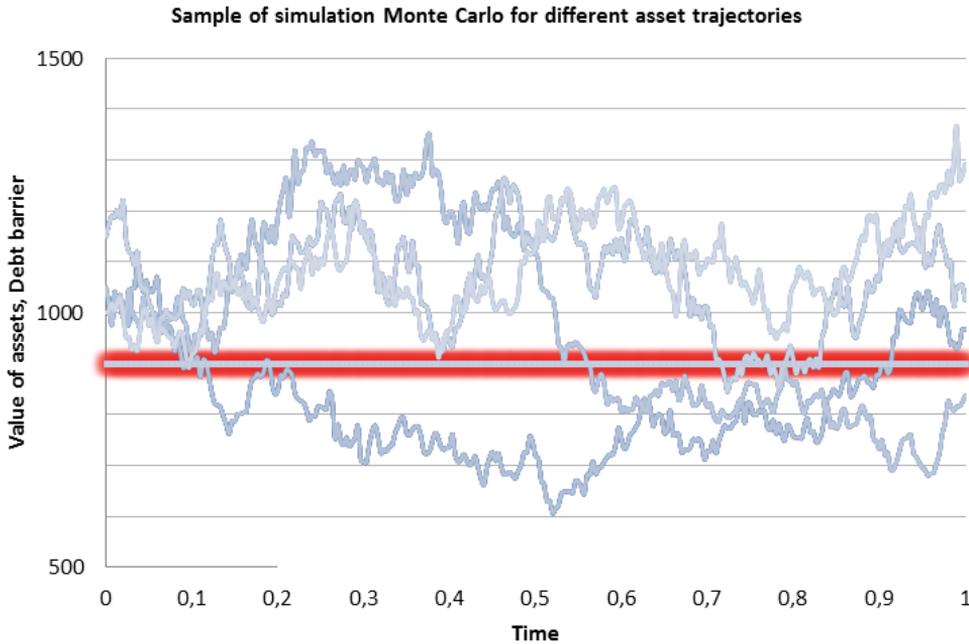Sample of simulation Monte Carlo for different asset trajectories

Fig. 1: Draw from Monte Carlo asset trajectories for company assets of one business entity
(trajectories ending under the red barrier in time 1 signs default scenario).

process), and $\sigma > 0$ is the volatility (standard deviation), V holds for firm asset value. The related log-returns of asset values are then

$$\log V_t - \log V_0 = \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t, \quad (2)$$

which follow a Normal distribution, $N\left((\mu - \sigma^2/2)\, t, \sigma^2 t\right)$. Thus V has a Lognormal distribution.

As stated in Klepáč (2014), Merton model is used for estimate the survival probability of chosen entity. In this model the company assts include equities and liabilities. Asset value of the entity $V = \{V_t, 0 < t < T\}$ is the sum of the equity value, $E = \{E_t, 0 < t < T\}$ and the value of a zero-coupon bond $z^t = \{z^t, 0 < t < T\}$ with maturity T and face value L: $V_t = E_t + z_t^T$. Default occurs if, at maturity, the asset value is not enough to pay back the face value L, see Fig. 1 for a better recognizance.

In this case the beholders take control of the firm and the shareholders do not receive anything. If at point of maturity $V_t > L$, default does not occur and the shareholders receive $V_t - L$. These assumptions allow us to treat the company equity as a European call option with inducted pay-off structure

$$
\begin{aligned}
E_t &= \max\left(V_t - L, 0\right) = \quad (3) \\
&= \begin{cases} V_t - L, & \text{if } V_t > L \text{ (no default)}, \\ 0, & \text{if } V_t < L \text{ (default)}. \end{cases}
\end{aligned}
$$

As mentioned above the equity can be seen as a European call option, thus we can use standard Black–Scholes partial differential equation as modelling tool

$$S_t = V_t N(d^+) - e^{-r(T-t)} K N(d^-), \quad (4)$$

where $N(d^{\pm})$ are distribution functions of a Standard Normal random variable, r is risk-free rate, K is nominal value of debt. We could calculate values for $d^{\pm}$ as probability metrics, that asset path trajectories would end under default debt level

$$d^{\pm} = \frac{\ln \frac{A_t}{K} + \left(r \pm \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{T-t}}. \quad (5)$$

At selected time $0 < t < T$, the conditional probability that no default will occur in (t, T) corresponds to the probability of finishing in

the money for the virtual call option held by shareholders. From further investigations we know that the survival probability for entity equals to $N(d^-)$.

Part of the empirical studies, e.g. Leland and Toft (1996), verified that Merton's model or common structural models underestimate the prices of credit derivatives, thus also the default risk compared with empirical data in short-term period. Hillegeist et al. (2002) state, that results obtained via a Merton-based model provide up to 14 times higher information value

(statistically) when a bankruptcy is determined than both Altman *Z*-score and Ohlson *O*-score – these scores can offer mostly additional information. These results were obtained in developed markets from the extensive files of data of several hundred or thousands companies.

In the conditions of the Czech Republic, similar studies are lately scarce, similar methodology is used in Míšek (2006). This scarcity is due to the low level of financial market development and low number of traded nonfinancial companies.

3   METHODOLOGY AND DATA

We deal with annual accounting data from yearly reports and Patria Online databases for market data (time series of closing stock prices, transformed into log returns). Data consists of four traded nonfinancial companies from Prague Stock exchange (PSE). Due to the potential use of presented methodology and its generalization is not necessary to fully mention the companies.

3.1   Estimation of default probabilities by Merton model

The main aim of this contribution is to test the possibility and compare results of default probabilities of the use of the basic Merton and enhanced Merton models of credit risk on the data of above stated companies. In this context, the risk of credit situation in three years (from 2011 to 2014), will be estimated and the results of obtained risk indicator (default probability) will be evaluated. The analysis proceeds from the medium term risk of credit situation, for the 4 year ahead estimate. With regard to the permanent financial market off, it will be considered if these models can point out the quality and changes in companies' financial structure – whether there have been more distinct changes and whether the values of estimates match the values of issued bonds at least approximately. At first instance we have to derive parameters:

- volatility of equity from stock daily returns;

- market value of equity which equals to market capitalization;

- risk-free interest rate from EU;

- time of debt settlement;

- liabilities – there exists many possible methods for estimate, but we use KMV's methodology. So the default level at maturity equals short-term + one half of long-term debts.

Volatility estimation based on combined conditional mean and variance leverage model ARMA(1,1)-GARCH(1,1)-GJR with Student-*t* distribution of innovation process was performed for time period from January 2007 to December 2011 are presented, then we took estimated equity volatility for calculation in Merton model, see below. This specific model setting performed best from in in-sample testing in our previous work, e.g. see Klepáč and Hampel (2015) for more details about techniques for models selections. The estimation of time-varying equity volatility and solving of simultaneous system (to get asset volatility and its market value) of equations were performed in software (SW) package R 3.1.1. (2015) and SW Matlab 2014b (MathWorks, 2014). Specifically, we followed these steps:

- Calculation of the company's market asset value, market value of debt, quantification of the theoretical default level of bankruptcy

when we use KMV methodology for default level selection.

- The volatility of asset value and market asset value is determined via calculation of system of non-linear equations (full market capitalization, historical yearly volatility mean estimate, the risk-free interest, which depend on the face value of debt, equity capital value on financial market rate – from EU countries yield curve), and other possible factors depending on the model complexity. For exact solutions of chosen factors see Merton (1973) or Míšek (2006).

- Calculation of the probability of default for chosen time periods and model settings.

3.2   Copulas and D-Vine copulas

Copulas were firstly introduced in mathematical context by Sklar (1959) through his famous theorem. Any multivariate joint distribution can be written in terms of univariate marginal distribution functions and a copula describes the dependence structure between the variables. Continuous development of the copula theory supports us with solutions for elliptical and non-elliptical distributions, see Joe (1996) for mathematical reference about these copulas. For estimations we use maximum-likelihood estimation (MLE) – where particular copulas have one or two parameters. For high dimensionality treatment many authors proposed pair-copula constructions (PCCs) class which structures partly shown in Fig. 2.

Among the specific types within this class are Vine copulas, which are a flexible class of n-dimensional dependency models when we use bivariate copulas as building blocks. Due to Aas et al. (2009) who described statistical inference techniques, we can create multi-tier structure (according to dependence intensity between variables) between one central variable (i.e. market index) and underlying variables (companies in this index). D-Vines in contrast to its "sibling" offer another view: we propose modelling of the inner structure without selection of one explicit dependence driver.

3.3   Construction of Merton-D-Vine copula model

The theoretical assumption is that the credit quality of a company, which is expressed via the trend of development of the company's assets, is developing together with the development of its shares in a certain way. The next assumption is the possibility that the development of the company's assets is making a progress within the framework of an exact system, where the relationship between the returns within the framework of defined groups can be measured.

Based on these assumptions, a prediction model can be defined which projects the dependences among companies on the financial market into the development of their assets. The aim of the model suggested here is to show the probability of default which can be used as the variable in default and classification models, because of illustration of dependence between data on stock market.

Specifically, the behaviour of the basic Merton model is simulated, based on the numerical solution of the model of European call option with parameters which are routinely used here while estimating default. The difference is that the value of assets is directed by Brown movement whose component, which usually matches the Wiener process with normal distribution, is replaced by D-Vine copula component with Student-t or Normal distribution. Thus the above mentioned relation as in (2) is used

$$dV_t = \mu V_t dt + \sigma V_t dW_t, \qquad (6)$$

which is generalized to the form of univariate asset trajectory model by D-Vine copula innovations

$$dV_t = \mu V_t dt + \sigma V_t dt \,(\text{D-Vine copula}_t), \quad (7)$$

and generalization for i companies

$$dV_t^i = \mu^i V_t dt + \sigma_i V_t dt \,(\text{D-Vine copula}_t^i), \quad (8)$$

where σ_i signs annualized asset volatility, which is a constant until debt settlement. D-Vine copula is a random draw from normalized D-Vine copula distribution, for capturing returns correlations. Mean of assets value process is
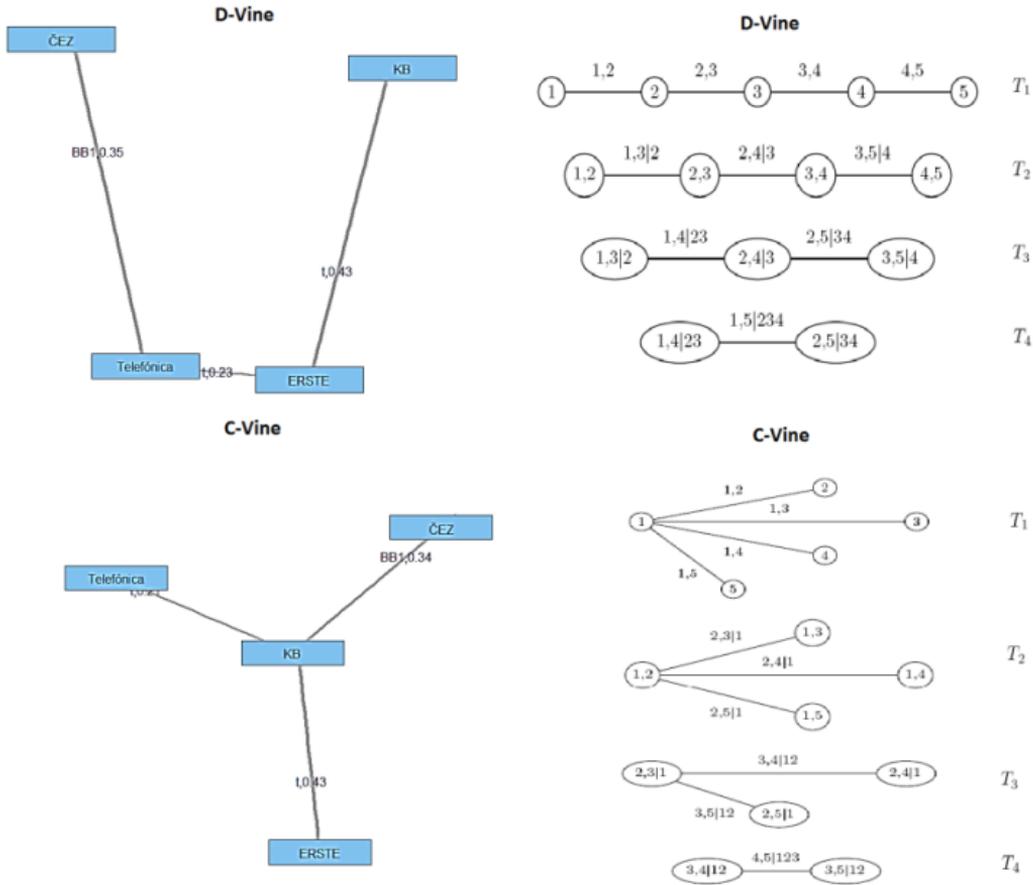
Fig. 2: Examples of D-Vine tree (right upper panel) and C-Vine tree (right panel bellow) specification on five variables, on the left panels are samples from four-dimensional D-Vine (left upper panel) and C-Vine copula (left down panel). The both left panels show example of the first structure. Based on Klepáč and Hampel (2015).

driven by process described by Campbell et al. (2008), where $\mu = r + 0.06$, and r is risk free rate.

The diversion from the standard Merton process lies in the fact that the innovations of the process contain the inner correlation structure and will be generated by non-normal division. Technically, it is inverse transformation performed through the *tcdf* algorithm in the Matlab 2014b, where the input is formed with the data from the D-Vine copula simulation, which is governed by uniform distribution on the interval $[0; 1]$. In the case of numerical option pricing we use steps described by Goddard Consulting (2011):

1. The calculation of the future development of the value of underlying asset – the output is formed by hundreds of trajectories of development, based on the determined function of the development of the asset, when the development is divided into small discrete intervals.

2. Calculation of terminal values of options for each of the potential trajectories.

3. Calculation of the average of all terminal values of options and their discounting in order to achieve the present value.

Thus the aim is to determine the frequency of intersections of the default barrier in the time of bond maturity, which can be – according to the

option theory – understood as the probability that the company will default at the concrete moment of time. That is why only the first steps from the above mentioned are used.

To estimate D-Vine copula on market data we should proceed in steps provided by Aas et al. (2009). At the first preparatory stage we filter (fit) raw data with ARMA(1,1) model, standardize residuals by GARCH (1,1) volatility – with the best fitting univariate models, in our case by GJR. Then we transform the residuals into uniformly distributed data from [0; 1] interval, we use algorithm *tcdf* in Matlab for this purpose. With the data in this form we could proceed according to Aas et al. (2009) and conduct:

- Structure selection to assess the intensity and structure of dependence.
- Copula selection for the most appropriate fit of the tail dependence characteristics with Vuong-Clark test.
- Estimation of copula parameters with maximum-likelihood estimation (MLE), when we use copulas with one or two parameters.
- Model evaluation by Vuong test and subsequently by information criteria.
- Simulation from D-Vine copulas to get at least 10,000 times n-asset of uniformly distributed numbers.
- Simulation based credit risk evaluation with Merton-D-Vine copula model.

4 RESULTS

In Tab. 1 there are presented basic returns statistics from which is obvious that median and mean values were around zero.

On Fig. 3 are visualized time-varying standard deviations (volatility) fits for capturing higher oscillations in time. According to previous results, see Klepáč and Hampel (2015), we preferred fits based on ARMA(1,1)-GARCH(1,1)-GJR which offers better quality in terms of information criteria due to the ability to measure non-linear data pattern in opposition to standard GARCH model. The highest daily oscillations lasted from 2008 to 2009 depending to global financial market crisis, see Fig. 3. Monitoring other higher daily returns we see breaks in 2011 and 2012 and in the summer of 2015.

4.1 Estimation of D-Vine copula models

For the Vine copula function estimates, first the dependence between magnitudes and the character of their distribution must be clarified and specified in detail, although not fully reported. From exploratory data analysis and the statistical inference procedures we have

drawn more specific conclusions about copula type which we have further used. More detailed testing provides us with the families of the copula models, which expresses this structure of dependence in the most appropriate way: we should deal with BB1, BB7 or Student-*t* copula families. After that we let the D-Vine algorithm separate the residuals into 3 trees. We recognized from "upper" into "bottom level" dependence intensity for these copula pairs (sign *t* denotes Student-*t* distribution, BB7 and BB1 are other copula types, number quantifies Kendall's tau correlation coefficient), see Fig. 4 where at first simple dependence structure is used. Therefor the more complex structure with weighted densities between return streams and lesser correlation is visualized in the case of Tree 2 and 3. The highest dependence is between the Companies 1 and 2, so the leading variable – dependence driver is Company 1, what manifests into the second dependence tier in Tree 2.

From this multivariate copula, visualized by the branch structure, we generate simulated uniformly distributed numbers from interval [0; 1] for transformation with inverse cumulative distributions from simulated probabilities.

Tab. 1: Basic statistics for returns data set

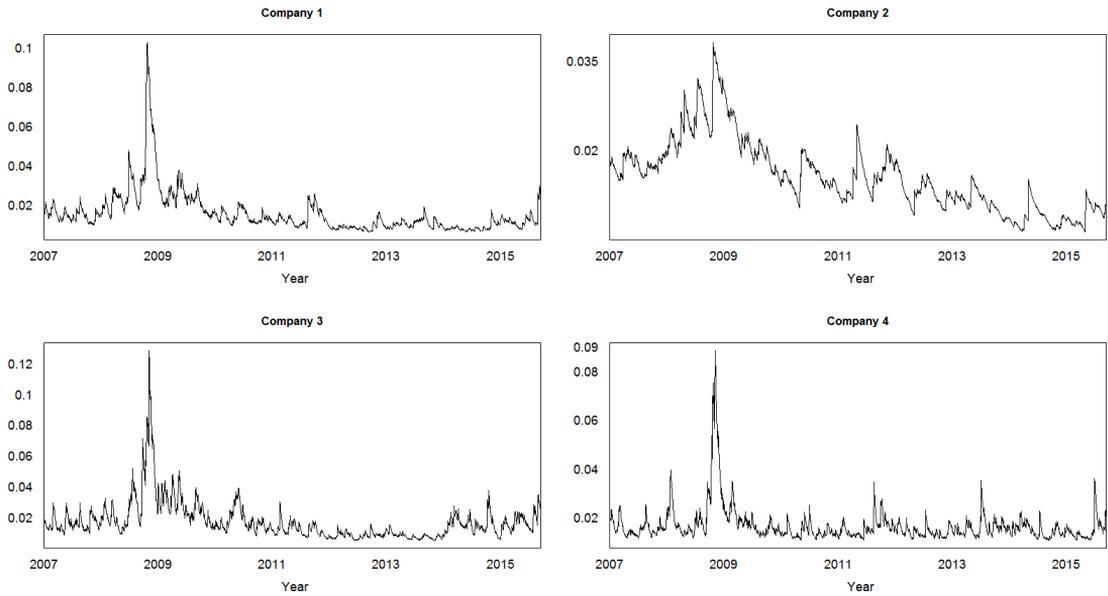| Statistics	Company 1	Company 2	Company 3	Company 4
Minimum	−0.2337	−0.1240	−0.1670	−0.1530
1st Quartile	−0.0066	−0.0066	−0.0077	−0.0090
Median	0.0000	0.0000	0.0000	0.0001
Mean	0.0001	0.0000	−0.0002	−0.0002
3rd Quartile	0.0070	0.0070	0.0081	0.0092
Maximum	0.1300	0.1480	0.2640	0.1951

Fig. 3: Daily volatility estimated for analyzed companies – with ARMA(1,1)-GARCH(1,1)-GJR model
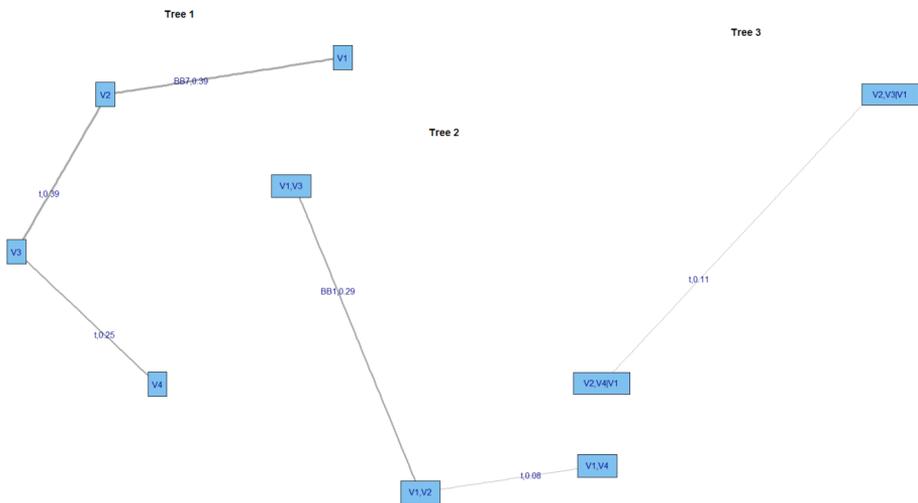

Fig. 4: D-Vine copula trees (V1–V4 sign companies, first expression is copula family, second evaluates Kendall tau value)

Tab. 2: Inputs for risk computation (in CZK) for 2007–2011 and 2011–2014 risk prediction

Input values	Company 1	Company 2	Company 3	Company 4
Market capitalization (Equity)	4,365,506,200	28,002,937,200	36,085,618,036	$4.30392 \cdot 10^{11}$
Debt barrier at maturity	1,730,643,600	6,564,000,000	$1.85 \cdot 10^{10}$	$1.01071 \cdot 10^{11}$
Annualized stock mean volatility	33.52%	33.27%	38.13%	29.54%
Market value of debt	1,640,493,800	6,227,062,800	$1.746938 \cdot 10^{10}$	$9.5908 \cdot 10^{10}$
Market value of assets	$6.006 \cdot 10^9$	$3.423 \cdot 10^{10}$	$5.355 \cdot 10^{10}$	$5.263 \cdot 10^{11}$
Annualized assets volatility	24.41%	27.27%	25.89%	24.16%
3-year EU27 risk-free rate	1.31%	1.31%	1.31%	1.31%

Tab. 3: Output for risk computation (in CZK) for 2007–2011 and 2011–2014 risk prediction

Calculation outputs	Company 1	Company 2	Company 3	Company 4
Expected Loss	0.0011	0.0003	0.0049	0.0001
Probability of Default – 4 year ahead	0.0079	0.0022	0.0291	0.0005
Distance to Default	2.9777	2.9999	2.6019	3.3850
Yield Spread (basis pts)	2.7855	0.7577	12.3032	0.1469
Probability of Default Merton-Copula	0.0781	0.0468	0	0

4.2 Estimation of default probabilities

Initial information resulting from calculation and used for risk estimation are offered in Tab. 2, where are input values of high importance: market value of assets and its volatility, Debt barrier at maturity, Equity value, EU27 3-year risk-free rate.

The resulting values of default probabilities for the 4 years ahead are visible in Tab. 3. Merton model usually underestimates realized level of risk for shorter maturities, which holds true in our case either, but adjusted Merton D-Vine copula model performs better – proposed probabilities based on Monte Carlo simulations are much higher, especially for Company 1 (which is most risky according to enhanced Merton model) and 2.

5 DISCUSSION AND CONCLUSIONS

Presented contribution concentrated on measurement of credit risk, probability of default in time in case of four traded nonfinancial companies from Prague market exchange. We analysed default probabilities of these companies and took in account the most renowned structural model of default – Merton model and proposed novel static Merton-D-vine copula model, which transforms market risk dependence to credit dependence with capturing nonlinear and high dimensional attributes.

Although Merton structural model usually underestimates realized level of risk for shorter maturities (like 1 or 3 years), but adjusted Merton model performs better – proposed

probabilities are much higher. Definitely, standard Brownian asset process in Merton model suffers with many weaknesses connected to Gaussian process property, that means we adjusted distribution of process too – instead of Gaussian distributions we propose Student-t distribution as a better choice because of its heavier tails. For a better understanding it is useful to compare these results in the future with models based on accounting data or data mining methods with bankruptcy prediction in mind or set obtained default probabilities as predictors in comparison to standard financial ratios.

6 ACKNOWLEDGEMENTS

7 REFERENCES

Aas, K., Czado, C., Frigessi, A. et al. 2009. Pair-copula constructions of multiple dependence. *Insurance: Mathematics & Economics.*

Altman, E. I. 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, 23 (4), 589–609.

Beaver, W. 1966. Financial ratios predictors of failure. *Journal of Accounting Research*, 4, 71–111.

Campbell, J., Hilscher, J. and Szilagyi, J. 2008. In search of distress risk. *Journal of Finance*, 63, 2899–2939.

Goddard Consulting. 2011. *Option Pricing – Monte-Carlo Methods* [online]. Available at: http://www.goddardconsulting.ca/option-pricing-monte-carlo-index.html.

Hillegeist, S. A., Keating, E. K., Cram, D. P. and Lundstedt, K. G. 2002. Assessing the Probability of Bankruptcy. *SSRN Electronic Journal.* DOI: 10.2139/ssrn.307479.

Joe, H. 1996. Families of m-variate distributions with given margins and $m(m+1) = 2$ bivariate dependence parameters. In Rauchendorf, L., Schweizer, B. and Taylor, M. D. (Eds.). *Distributions with mixed marginals and related topics.*

Klepáč, V. and Hampel, D. 2015. Assessing Efficiency of D-Vine Copula ARMA-GARCH Method in Value at Risk Forecasting: Evidence from PSE Listed Companies. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 63 (4), 1287–1295. ISSN 1211-8516.

Klepáč, V. 2014. Assesing Probability of Default: Merton Model Approach. In *PEFnet 2014*. Brno: Mendel University in Brno, 76.

Leland, H. E. and Toft, K. B. 1996. Optimal Capital Structure, Endogenous Bankruptcy and the Term Structure of Credit Spreads. *Journal of Finance*, 51 (3), 987–1019.

Longstaff, F. A. and Schwartz, E. S. 1995. A Simple Approach to Valuing Risky Fixed and Floating Rate Debt. *Journal of Finance*, 50 (3), 789–819.

Market data by Patria Online. [online]. Available in paid version at: http://www.patria.cz/Stocks/default.aspx.

MathWorks. 2014. *MATLAB and Statistics Toolbox Release 2014b.*

Merton, R. 1973. On the Pricing of Corporate Debt: The Risk Structure of Interest Rate. *Journal of Finance*, 29 (2), 449–470.

Míšek, R. 2006. *Strukturální modely kreditního rizika.* (Ph.D. thesis.)

Ohlson, J. A. 1980. Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, 18 (1), 109–131.

R Development Core Team. 2015. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.

Schoutens, W. and Cariboni, J. 2009. *Lévy processes in credit risk.* Chichester, U. K.: John Wiley, 185 pp. ISBN 04-707-4306-9.

Sklar, A. 1959. Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8, 229–231.

Zmijewski, M. 1984. Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 59–86.

AUTHOR'S ADDRESS

Václav Klepáč, Department of Statistics and Operation Analysis, Faculty of Business and Economics, Mendel University in Brno, Zemědělská 1, 613 00 Brno, Czech Republic, e-mail: xklepac@mendelu.cz