Data Science.
Lectures. Weeks 3-4.
Forecasting the default probability without accounting data.
Прогнозирование вероятности дефолта без данных бухгалтерского учета.

Polina Kravets

9 октября 2022 г.

Contents

1 Introduction

1.1 A brief review of the Merton model (Merton-type models):

A very common assumption in Merton-type models (and Moody's KMV) is that the value A_t of the firm follows a geometric Brownian motion:

$$dA_t = \mu_A A_t dt + \sigma_A A_t dW_t \tag{1}$$

where W_t is a Wiener process, with drift and volatility coefficients μ_A and σ_A.

Then, we have:
Call option price:

$$E_t = A_t N(d_1) - \exp^{-rT} B_t N(d2) \tag{2}$$

$$d_1 = \frac{\log(\frac{A_t}{B_T}) + (\mu_A + 0,5\sigma_A^2)T}{\sigma_A \sqrt{T}} \tag{3}$$

$$d_2 = d_1 - \sigma_A \sqrt{T} \tag{4}$$

Standard deviation of the equity:

$$\sigma_E = \frac{A_t}{E_t} N(d_1)\sigma_A \tag{5}$$

We don't know the acid value time t and we don't know the standard deviation of the acid. $\Rightarrow$
Take (2) and (5) and solve them numerically for A_t and σ_A, while the implied probability of default (cumulative distribution function) is

$$P[A_T \leq B_T = N(-d_2)]$$

1.2 Related problems:

Structural models: They are extensions of the Merton's model, where Firm's capital is a residual concept and and the Firm's Value A_t is exogenous. However, this is not realistic today:

- Stocks belonging to workers cannot be used in a default;
- Stocks belonging to managers can be only partly used (usually not the salary part);
- If T is not the "end of the World", but simply a default, the law usually blocks the creditors in order to let the firm start again;
- A physical asset has a different value, according to the different source of financing: look at Fiat - GM negotiations, Opel ext.;
- As a consequence, the Firm's Value A_t is *endogenous*;

Reduced Form models (last 10 years): Pure data-fitting models where the firm's default probability is modelled as a Poisson process (idea similar to earthquakes forecasting in Geology).

2 A New Approach for Firm Value

2.1 An Unfeasible Approach

A more realistic framework: bivariate contigent claims: In order to solve the previous problems, we propose to resort to the theory of bivariate contingent claims. A *contingent claim* can be written in the general form as:

$$G(f(S_1(T), S_2(T); T)$$

where $G(\cdot)$ is a univariate pay-off function which identifies the derivative contract, $f(\cdot)$ is a bivariate function which describes how the 2 underlying securities determine the final cash-flows, S_i denotes the price of the i^{th} underlying security and T is the contract maturity.

By using this framework, we can express the *value of the firm* A_T as

$$A_T = G(E_T, B_T; T) = E_T + B_T$$

$$f(E_T, B_T; T) = I_{[(\mathrm{E}_T \geq 0), (0 \leq B_T \leq D)]}$$

where I is the indicator function.

The case of complete markets:

- The bivariate contingent claim can be exactly replicated and its price is uniquely determined;
- There is a unique *risk-neutral probability distribution* $Q(E, B|\mathcal{F}_t)$, with density function denoted by $q(E, B|\mathcal{F}_t)$, which represent the pricing kernel of the economy;

$$A_t = g(E_t; B_t; t) = P(t, T) \int_0^\infty \int_0^D G(E_T, B_T; T) q(E_T, B_T|\mathcal{F}_t)) dE_T dB_T \tag{8}$$

where D is the bond face value, while $P(t, T)$ is the risk-free discount factor, which we assume deterministic or independent of E_T and B_T , for sake of simplicity.

⇒ By using the Sklar's Theorem (1959), we can write the joint pricing kernel $q(E, B|\mathcal{F}_t)$ as a product between and the copula[1] and marginal densities:

$$q(E, B|\mathcal{F}_t) = c_{EB}(Q_E, Q_B|\mathcal{F}_t) \cdot q_E(Q_E|\mathcal{F}_t) \cdot q_B(Q_B|\mathcal{F}_t) \tag{9}$$

where c_{EB} is the density associated to the copula function.

⇒ The firm's value price $g(E_t, B_t; t)$ can be therefore expressed as,

$$A_t = g(E_t, B_t; t) = P(t, T) \int_0^\infty \int_0^D G(E_T, B_T; T) c_{EB}(Q_E, Q_B|\mathcal{F}_t) \cdot q_E(Q_E|\mathcal{F}_t) \cdot q_B(Q_B|\mathcal{F}_t) dE_T dB_T \tag{10}$$

Why is this approach infeasible (for now) ?

- The OTC derivative market on corporate bonds is very limited and data are not available;
- Usually, only part of a firm's debt is quoted, and in many cases can be not quoted at all.

2.2 A Feasible Approach

When bonds are traded and liquid, a more realistic approach is to consider the joint probability distribution for stocks and bonds, and discount the bivariate contingent claim with a risky discount factor.

By using the Sklar's Theorem, the new pricing function A'_t is:

$$A'_t = P_i(t, T) \int_0^\infty \int_0^D \int_0^\infty G(E_T, B_T; T) c_{E,B,i}(F_E, F_B, F_i|\mathcal{F}_t) f_E(E|\mathcal{F}_t) f_B(B|\mathcal{F}_t) f_i(i|\mathcal{F}_t) dE_T dB_T di_T \tag{11}$$

where $P_i(t, T)$ is the risky discount factor, i_T is the risky interest rate, $c_{E,B,i}$, f_E, f_B, f_i are the risky copula density, the stock, bond, interest rate marginal density functions, respectively.

⇒ Whenever the bond issues are *illiquid or are not traded* at all, we can express the bond price as a function of the risky interest rate, instead.

In such a situation, the previous expression (11) can be suitably modified:

$$A''_t = P_i(t, T) \int_0^\infty \int_0^\infty G(E_T, B_T(i_T); T) c_{E,i}(F_E, F_i|\mathcal{F}_t) f_E(E|\mathcal{F}_t) f_i(i|\mathcal{F}_t) dE_T di_T \tag{12}$$

[1]AN - copula is an approach that allows you to model more general multiplier distribution function

where $c_{E,i}$ is the copula density between stock and risky interest rates.

⇒ The previous formula can be further simplified by considering that B_T is known at time t as well as $P_i(t,T)$.

As a consequence, we need the stock price distribution, only:

$$A_t''' = P_i(t,T)\int_0^{\infty} G(E_T, B_T; T) f_E(E|\mathcal{F}_t) dE_T \tag{13}$$

The previous firm pricing function (13) can be approximated by Monte Carlo methods as follows:

$$\tilde{A}_t = P_i(t,T)\frac{1}{N}\sum_{i=1}^{N} G(\tilde{E}_{i,T}, B_T; T)$$

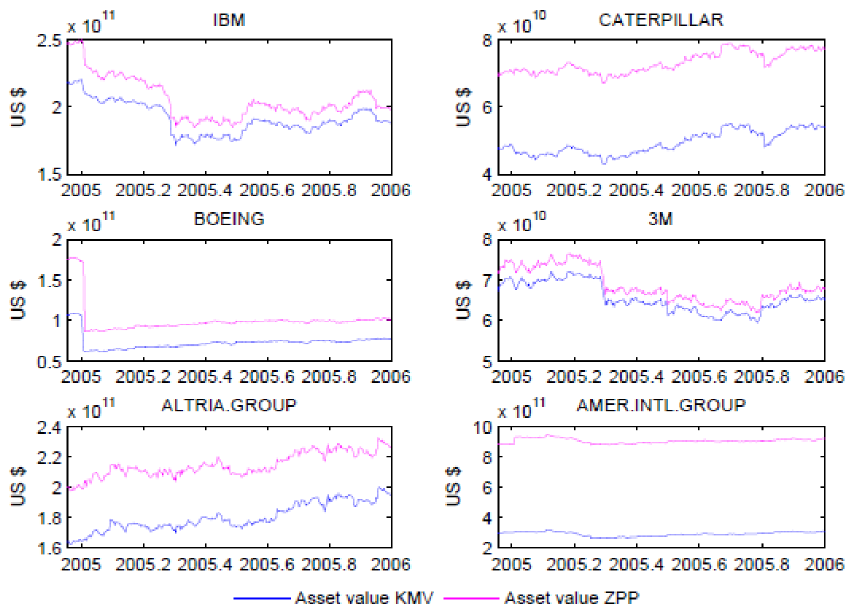

Fig. 1: Firm Value - stocks DOW30

3 A New Approach for Default Probability Estimation: The Zero-Price-Probability

However, if we theoretically allow the domain of ET to range from -∞ to +∞ and consider prices in levels instead of log prices, this interesting result follows directly:

Proposition 1 [Fantazzini, DeGiuli, Maggi (2008)]: The Default Probability is given by $P(E_T < 0)$ or $P(P_T < 0)$, given that $E_T = SP_T$, where P_T is the stock price at time T and S is the no. of shares.

Since $P_T = max(P_T, 0)$ is a particular truncated variable, the Default Probability is the probability that PT goes below the truncation level of zero, or simply the Zero-Price Probability (ZPP).

If we consider the following two quantities:

$$\begin{cases} E_T = A_T - B_T \\ E'_T = A_T = (A_T - B_T) + B_T = E_T + B_T \end{cases}$$

we can easily see that the meanings and signs of E_T and E'_T can be completely different according to the situation faced by the firm:

Table 3: **Financial Meaning and Signs of E_T and E'_T**

	$E_T = A_T - B_T$	$E'_T = A_T$
OPERATIVE	Equity belonging to shareholders (+)	Asset value (+)
DEFAULTED	Loss given default for Debtholders (−)	Equity belonging to Debtholders (+)

⇒ Therefore, we can estimate the *Distant to Default (D.D.)* simply by using E_T, instead of Merton's formula $[ln(A_T) - ln(B_T) - T \cdot (\mu_E - \sigma_E/2)]/\sigma_A$, and the default probability by $P(E_T < 0)$.

If we are at time t and we want to compute the (implicit) probability at a given time $t+T$ that the stock price will cross the truncation level of zero, i.e. $p(P_t + T < 0)$, then

1. Consider a generic conditional model for the *differences of prices levels* $X_t = P_t - P_{t-1}$, *without the log-transformation*:
$$X_t = E[X_t|\mathfrak{S}_{t-1}] + \varepsilon_t \tag{15}$$
$$\varepsilon_t = H_t^{\frac{1}{2}} \eta_t, \eta_t \sim i.i.d(0,1) \tag{16}$$
where $H_t^{\frac{1}{2}}$ is the conditional standard deviation, while $\mathfrak{S}_t$ is the information set available at time t.
2. Simulate a high number N of price trajectories up to time $t+T$, using the estimated time series model
3. The default probability is simply the number of times n out of N when the price touched or crossed $P_T = 0$ along the simulated trajectory.

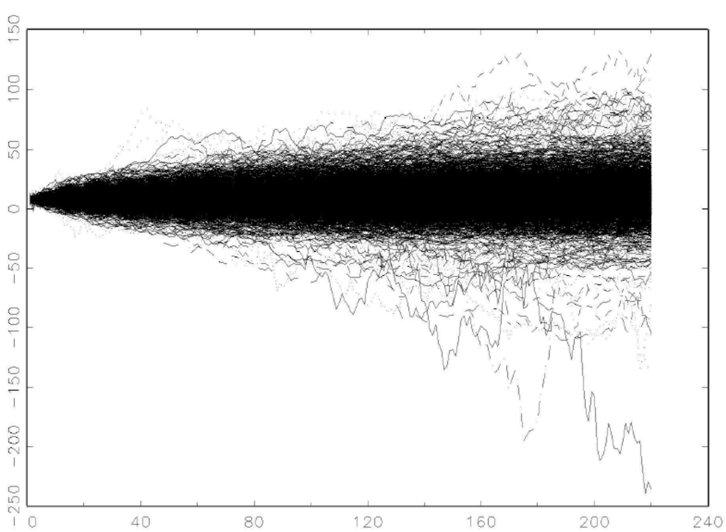

Fig. 2: Example: 5000 simulated price trajectories, 1-year ahead, for a risky stock (ZPP $\sim$ 40%)

This method entails a number of important benefits:

- We only need the stock prices for estimating the default probability;
- We do not need any latent default barrier D, or firm's volatility σ_A, like in Merton style models;

- We can estimate the default probability for any given time horizon $t + T$;
- We can consider more realistic distributions than the log-normal;
- We can screen the default risk daily or even intra-daily. The ZPP can therefore be used as a tool for risk management;
- Given the face value of the debt B_T, we can compute the average loss given default for debtholders and therefore the average recovery rate.

4 Empirical Evidence: American and European Markets

Fantazzini, DeGiuli, Maggi (2008) estimate the 1-year ahead ZPP considering the last 200 / 1000 trading days for four famous defaulted stocks[a]:

1. **Cirio**: 24/09/1999 - 24/07/2003 (Last 1000 days). Second largest default in the Italian food sector (the first is Parmalat, see below);
2. **Enron**: 13/02/2001 - 03/12/2001 (Last 200 days). Largest default in American history;
3. **Parmalat**: 22/02/2000 - 22/12/2003 (Last 1000 days). Largest default in Italian history;
4. **Swissair**: 12/12/2000 - 03/10/2001 (Last 200 days). Largest default in European Airline Industry.

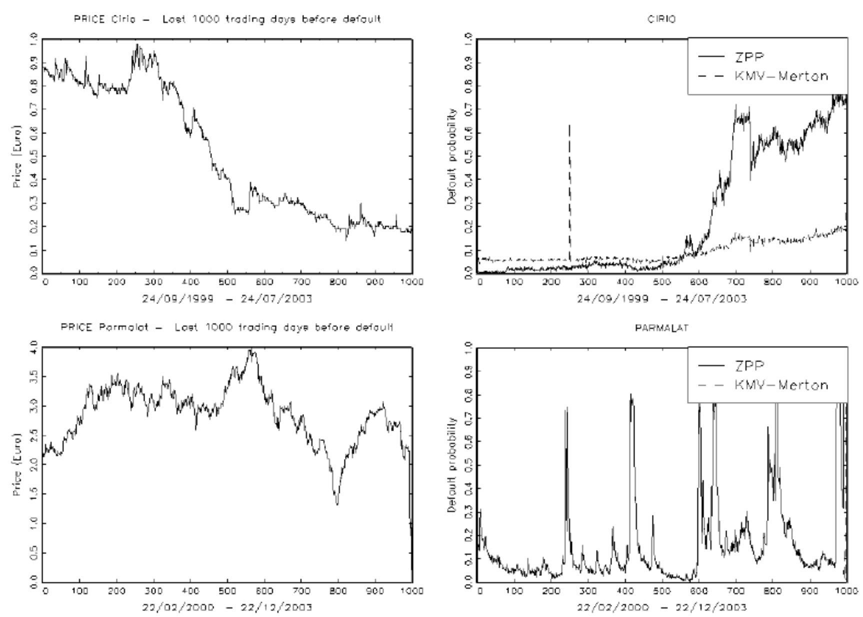

Fig. 3: KMV-Merton default probability and ZPP: CIRIO and PARMALAT

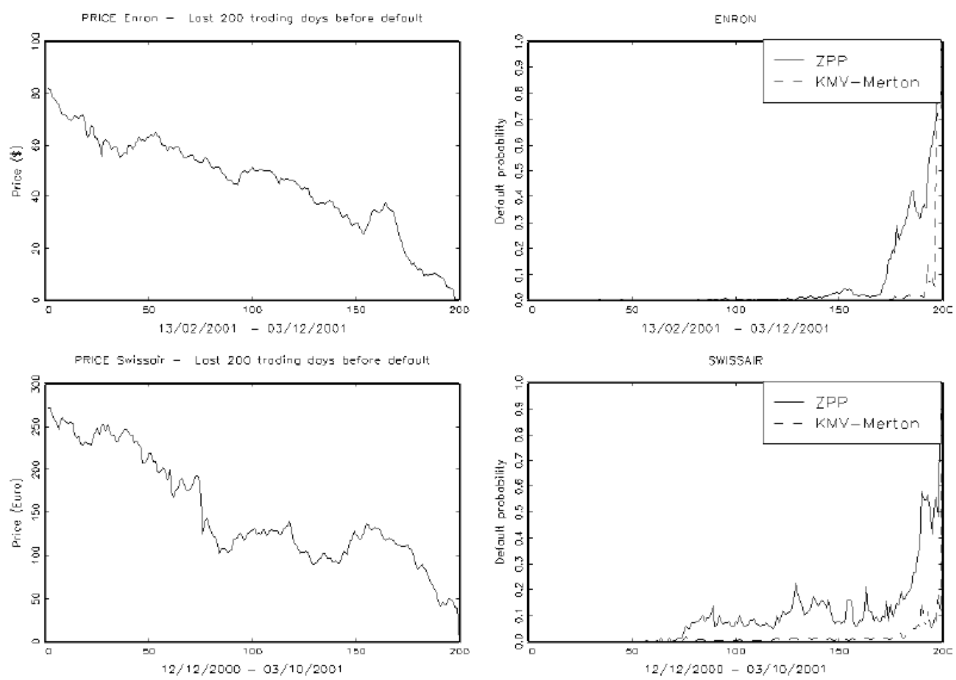

Fig. 4: KMV-Merton default probability and ZPP: ENRON and SWISSAIR

- The KMV-Merton model shows numerical instability due to abrupt changes in the debt values at the end of the year;
- The log-normal is not an appropriate distribution for price dynamics ⇒ tail underestimation;
- Debt values reported in the certified balance sheets are underestimated:
  ⇒ to “window dress“ the financial health of the company, in the best case (for instance, Swissair);
  ⇒ to hide financial fraud, in the worst case (Cirio, Enron, Parmalat).

We can gauge the precision of the estimated ZPPs by Monte Carlo methods:

1. *Draw a* $1T$ *vector of standardized innovations* η from the considered marginal density (for example Student’s T);
2. *Create an artificial history for the random variable* by replacing all parameters with their estimated counterparts, together with the standardized innovations η_t drawn in the previous step, which have to be rescaled by the the square roots of the variances $\sqrt{h_t}$;
3. *Estimate an AR(p)- GARCH model*, using the data from the artificial history;
4. *Calculate a Monte Carlo estimate of the ZPP* using the previous estimates performed on the artificial history;
5. *Repeat the above four steps for a large number of times*, in order to get a numerical approximation to the distribution of the ZPP.

This distribution forms the basis for computing a bounded kernel density.

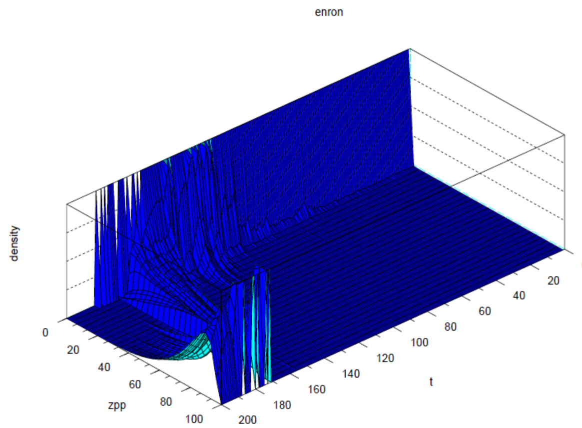

Fig. 5: Bounded kernel density for ENRON's Z.P.P.(last 200 trading days)

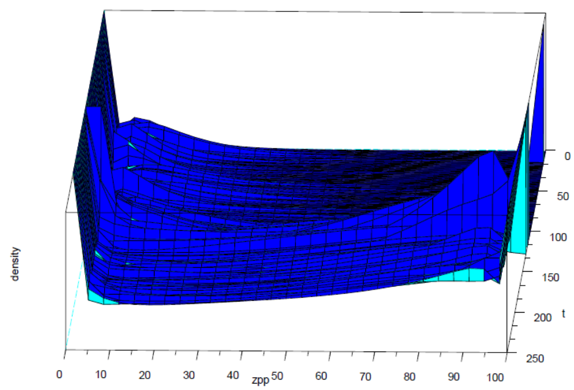

Fig. 6: Bounded kernel density for SWISSAIR's Z.P.P.(last 200 trading days)

5 Empirical Evidence: Russian Markets

Fantazzini (2009) analyzed the daily data of the **five most traded Russian stocks**: Gazprom, Lukoil, Norilsk Nickel, Sberbank and United Energy in the time sample 20/04/2006-23/04/2008.

⇒ He tested for unit roots in the financial variables under scrutiny by using the Dickey-Fuller Test with GLS Detrending (DF-GLS) by Elliott et al. (1996) and the test by Kwiatkowski, Phillips, Schmidt and Shin (1992), which is based on the null of covariance stationarity rather than integratedness.

	DF-GLS		**KPSS**	
Stock	**Levels**	**First differences**	**Levels**	**First differences**
GAZPROM	0.048	-26.170 (**)	4.479 (**)	0.168
LUKOIL	-0.615	-43.822 (**)	4.647 (**)	0.049
NORILSK NICKEL	0.595	-46.843 (**)	4.144 (**)	0.245
SBERBANK	-0.196	-39.150 (**)	4.135 (**)	0.229
UNITED ENERGY	-0.779	-41.649 (**)	4.162 (**)	0.169

⇒ Goodness-of-fit tests of the AR(1)-T-GARCH(1,1) models with Student's t errors, employed for the conditional marginal distributions:

- **Ljung-Box tests** on the standardized residuals in levels $\hat{\eta}_t$and squares $\hat{\eta}_t^2$;
- **Kolmogorov-Smirnov test** for density specification;
- **Hit test** by Granger et al. (2006), to test jointly for the adequacy of the dynamics and the density specifications in the marginal distribution models, where the null hypothesis is that the density model is well specified.

Stock	Ljung-Box(25) η_t	Ljung-Box(25) η_t^2	Kolmogorov--Smirnov	Joint Hit Test
GAZPROM	0.281	0.066	0.021	0.075
LUKOIL	0.334	1.000	0.604	0.032
NORILSK NICKEL	0.734	1.000	0.353	0.081
SBERBANK	0.632	0.987	0.107	0.056
UNITED ENERGY	0.854	0.331	0.282	0.074

- The default probabilities estimated with the ZPP are again higher than the ones obtained by using Merton's model;
- The KMV-Merton model shows some numerical instability problems with noisy data and from the jumps in the debt values at book closure dates at the end of the year.

⇒ More importantly **POLITICAL RISK** is not accounted for by the KMV-Merton model, while the ZPP can do that.

For example, look at the Gazprom's PD in the months before the Gazprom CEO was nominated as candidate for the Russian Presidency on the 11/12/2007:

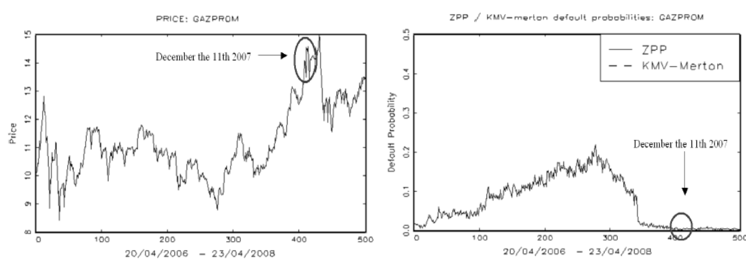

Fig. 7: KMV-Merton vs ZPP: GAZPROM

⇒ Let now consider **Yukos**, the largest default in Russian history.

⇒ We again employ AR(1)-T-GARCH(1,1) models with a Student's T distribution (Unit root tests and goodness-of-fit tests are reported below):

	DF-GLS		KPSS	
Stock	**Levels**	**First differences**	**Levels**	**First differences**
YUKOS	-0.238	-28.006 (**)	3.672 (**)	0.021

Stock	Ljung-Box(25) η_t	Ljung-Box(25) η_t^2	Kolmogorov--Smirnov	Joint Hit Test
YUKOS	0.263	0.405	0.556	0.237

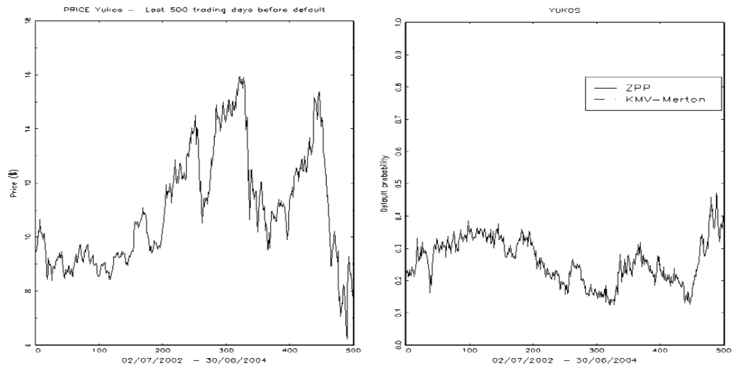

Fig. 8: KMV-Merton default probability and ZPP: YUKOS

⇒ The market seemed to price the difficulties regarding this company already a couple of months in advance the arrest of Yukos CEO in October 2003 (observation 332 in the two plots).

⇒ The estimated PD was already ranging between 15% and 30%. Again, the KMV-Merton model is unable to take the political risk into account.

6 Empirical Evidence: Global Financial Crisis

Fantazzini, Kudrov and Zlotnik (2010) **examined the development of credit risk** in the last two years (2007-2008), with particular reference to the Russian banking sector.

They analyze *four single banks* (one for Russia, one for USA, one for Italy and one for UK), that represent important cases due to their dimension and/or financial history:

- Sberbank (Russia)
- Citigroup (USA)
- Unicredit (Italy)
- Royal Bank of Scotland (UK)

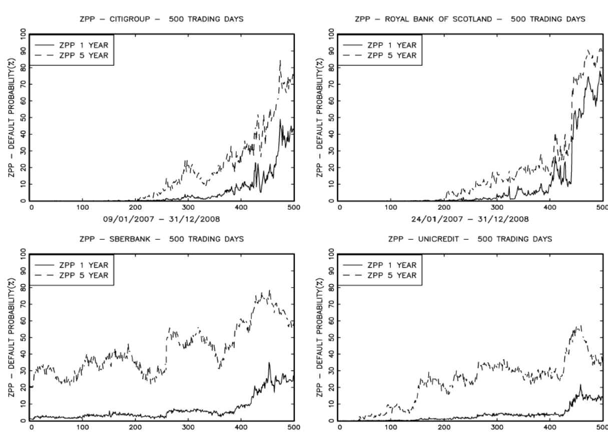

Fig. 9: Estimated Default Probability by using the ZPP: Citigroup, RBS, Sberbank and Unicredit

⇒ Second bailout for RBS in January-February 2009.

⇒ Citigroup split in two in January 2009 and a new capital infusion may be required after the recent "Stress tests" (to be made public on 04/05/2009, but partially leaked to the press).


TIMESONLINE

NEWS COMMENT BUSINESS MONEY SPORT LIFE & STYLE TRAVEL DRIV

MARKETS ECONOMICS INDUSTRY SECTORS COLUMNISTS MOVERS & SHAKERS LAW

Where am I? Home Business Industry Sectors Banking & Finance

From Times Online

January 19, 2009

RBS shares dive 70% on mounting debt fears


FT.com COMPANIES Financial Services

FINANCIAL TIMES

FT Home > Companies > Financials > Financial Services

Front page World Companies Energy Industrials Transport Retail & Consumer Health

Citigroup splits into two as it loses $8.3bn

By Alan Rappeport in New York

Published: January 16 2009 12:38 | Last updated: January 16 2009 15:50

Citigroup confirmed on Friday that it would split into two businesses after reporting a fourth-quarter loss of $8.29bn, or $1.72 a share.

⇒ As for Unicredit and Sberbank, even though their risks of default have increased after the financial turmoil in October 2008 , nevertheless these risks have stabilized since then, differently from the previous American and English banks.

As a confirmation of these insights, Fantazzini, Kudrov and Zlotnik (2010) *used a completely different methodology* based on Extreme Value Theory.

Robust estimation algorithm for Value at Risk:

1. For every consecutive 250 days (during the considered period of time) we compute the set of Hill's estimators ($\gamma(k)$) for the extremal index of the distribution function of negative returns:

   Suppose to have, within the considered period, m negative returns $X_1, ..., X_m$ and let $X_{(1)} \leq X_{(2)} \leq ... \leq X_{(m)}$ be their ordered statistics.

   Then the set of Hill's estimators ($\gamma(k)$) is defined as follows:

$$\gamma(k) = \frac{1}{k}\sum_{i=1}^{k}(lnX_{m-i+1} - lnX_{m-k}), 1 \leq k \leq m-1$$

   Consider the following model for the sequence of Hill's estimators ($\gamma(k)$):

$$\gamma(k) = \gamma + \beta_1 k + \varepsilon_k, k = 1, ..., k, \tag{17}$$

   where $E[\gamma(k)] = \gamma + \beta_1 k$, $V_{ar}(\varepsilon_k) = \sigma^2/k$ and γ is the true value of the extremal index for the distribution of negative returns.

   We can then estimate γ by using the method of weighted least squares with a weighting $k \times k$matrix W, that has $(\sqrt{1}, ..., \sqrt{k})$ on the main diagonal and zeroes elsewhere.

2. To estimate the excess level or Value a Risk x_p at the probability level p $(0 < p < 1)$ for the next day, we use the following estimator:

$$\hat{x}_p = \frac{\frac{r}{pn}^{\hat{\gamma}} - 1}{1 - 2^{-\hat{\gamma}}}(X_{(n-r)} - X_{(n-2r)}) + X_{(n-r)} \tag{18}$$

   where n is the number of negative returns, $r = [k/2]$ ([.]-integer part), $\hat{\gamma}$ is the estimator for the extremal index γ, $X_{(n-r)}$, $X_{(n-2r)}$ are $(n-r)$ - and $(n-2r)$ - ordered statistics of the absolute valued positive returns sequence $X_1, ..., X_n$, respectively.

⇒ The estimator of extremal index γ, used in the *first step* of the estimation algorithm for VaR was proposed by Huisman et al. (2001), where it is recommended to take $k = m/2$.

⇒ Instead of selecting an optimal threshold for the Hill's estimator of the extremal index, this approach allows to compute an optimal unbiased estimate of γ on the basis of the Hill's estimators set (with the thresholds $k = 1, ..., k$).

⇒ On the *second step* we use the consistent estimator of the excess level x_p proposed by Dekkers and De Haan (1989).

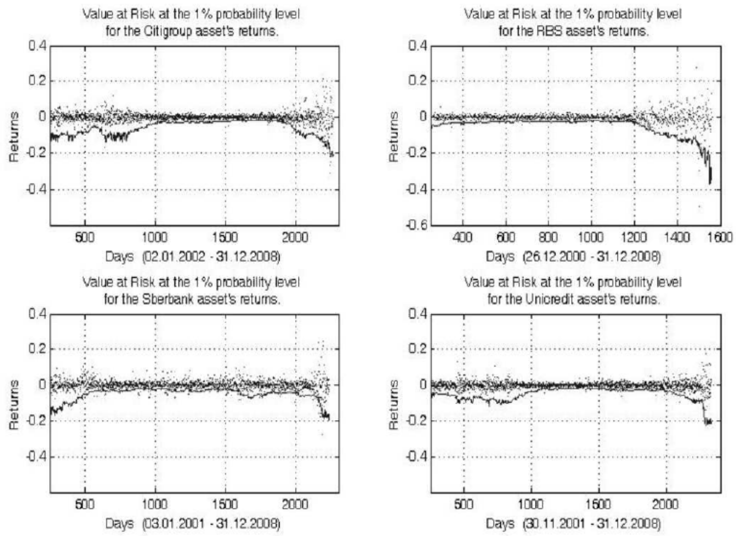

Fig. 10: Estimated Value at Risk at the 1% probability level: Citi-group, RBS, Sberbank and Unicredit.

The structure in Table 3 can be easily generalized to a **general sector**: instead of having the equity for a single firm, we can have the equity belonging to all shareholders of a specific sector, for example, the financial sector.

⇒ Therefore, by using a sector index instead of a single stock, the ZPP can also be used as an early warning system for systemic default of a general sector.

Fantazzini, Kudrov and Zlotnik (2010) consider the Russian RTS *Financial Index*, the American *Dow Jones Financial Index*, the English *FTSE Banking Index* and the Italian *MIBTEL Financial Index*.

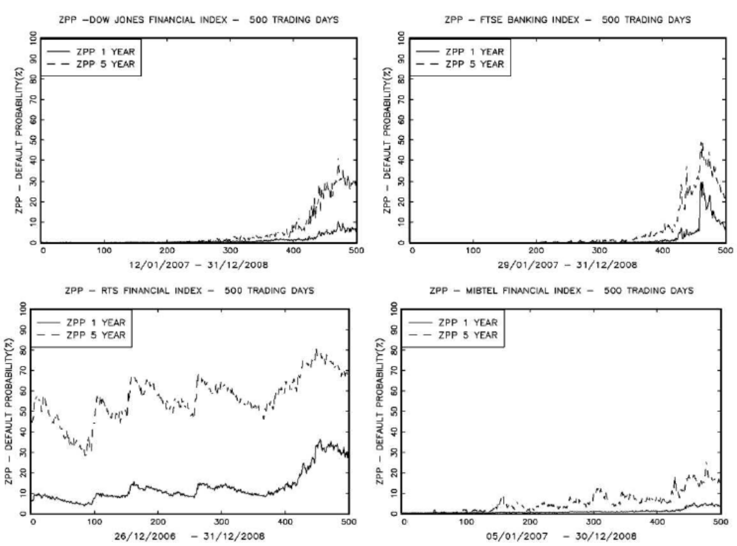

Fig. 11: Estimated Default Probability: American English, Russian and Italian financial sectors indexes

- The Russian financial index clearly shows a higher degree of riskiness than the other markets, that was quite high already at the beginning of 2007 and peaked in October 2008.
- However, this higher risk is mostly due to a higher country risk (Russia has a rating of BBB+), than the competing countries (US and UK have AAA, while Italy A+).

- Besides, the increases in the default probabilities for the American and English banking sectors in 2008 are very large and reflect the difficulties that they have experienced so far.

⇒ Interestingly, the Italian financial sector currently shows the smallest default probability (although it was still higher than the American and English ones till July 2008), thus confirming the smaller impact the subprime crisis has had on Italian banks

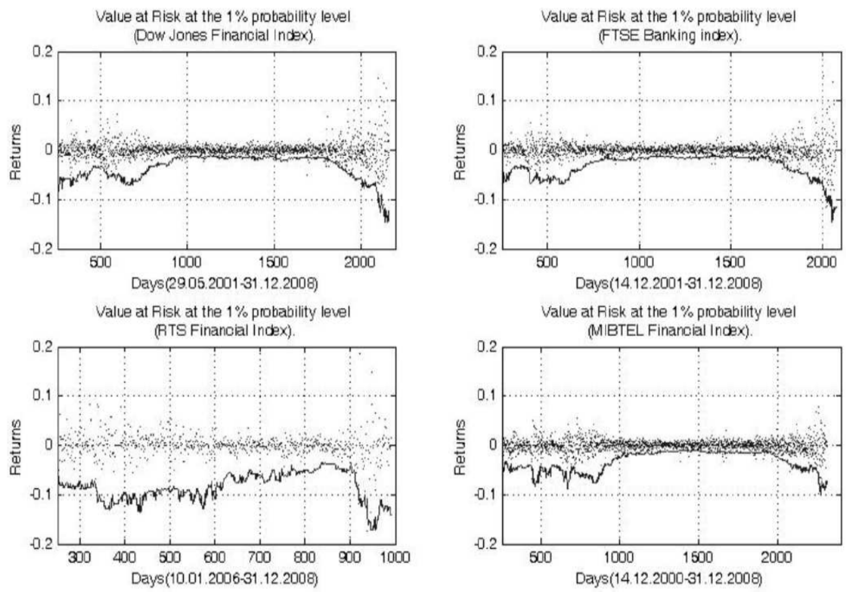

Fig. 12: Estimated Value at Risk at the 1% probability level: American, English, Russian and Italian financial sectors indexes

7 Conclusions

- *Contribution*: A new approach for firm value and default probability estimation was proposed.
- *Benefit (1)*: The Firm's Default Probability is a simple by-product and it does not even require the debt data, but stock prices, only.
- *Benefit (2)*: Lower and upper bounds of the firm's value and default probability can be provided by using bootstrap techniques.
- *Benefit (3)*: This new methodology is much more robust w.r.t financial frauds and political risk.
- *Avenue for future research*: Multivariate extension.
- *Avenue for future research (2)*: Perform a back-testing analysis with larger datasets