united nations
educational, scientific
and cultural
organization

international atomic
energy agency

the
abdus salam
international centre for theoretical physics

SMR/1235 - 5

Second School on the

MATHEMATICS OF ECONOMICS

(21 August - 1 September 2000)

Credit Risk
Theory, Measurement and Applications

M. Boldrin
(University of Minnesota)

strada costiera, 11 - 34014 trieste italy - tel. +39 040 2240111 fax +39 040 224163 - sci_info@ictp.trieste.it - www.ictp.trieste.it

Credit Risk
Theory, Measurement and Applications *

Michele Boldrin
University of Minnesota and CEPR, London
mboldrin@econ.umn.edu

May 1998; abridged August 2000

Contents

Introduction

Credit Risk is the most "ancient" and pervasive form of risk in the lending business.

Credit risk is also the most global or general equilibrium source of risk, as the creditworthiness of even a single line of credit depends on all kinds of factors: from the individual and idiosyncratic characteristics of the borrower, to movements in the national and international financial markets, to the behavior of other lending institutions, to business cycle movements both at the sectoral and aggregate level.

Strange enough, or maybe not so strangely, academic economists have started to develop formal tools to analyze, quantify and control CR. Very difficult tasks are often approached last.

i. From accounting instruments to statistical procedures.

ii. Adoption and application of contingent claims and no-arbitrage principles.

iii. Recognition that Assets and Liabilities Management (ALM) in a commercial bank is nothing but a special instance of portfolio management in front of uncertainty in market movements and idiosyncratic shocks.

iv. Applications of optimal portfolio theory (OPT).

v. Development of rigorous valuation models for borrowing instruments with uncertain future payment flows.

vi. Development of econometric and statistical techniques to measure and estimate such models.

Even if apparently obvious, it is crucial to understand that in facing CR we are not trying to eliminate it, but to manage it. Taking up risk is the intrinsic function of banks and financial market institutions: no risk, no (excess) return! Eliminating risk is trivial: just step out of the kitchen and lend only to sound and sovereign governments (or, do not lend at all). Managing risk is the true challenge.

1 Basic concepts

Credit risk refers to the event of a partial or total default on a promised payment, or sequence of them.

Loan losses reserves according to: (i) constant fraction of outstanding loans; (ii) peer equivalent behavior; (iii) historical loss records; (iv) management of fiscal burden; (v) static credit evaluation.

Credit evaluation methods. Most common credit scoring method is based upon linear multivariate scoring rules plus risk thresholds. Scoring rules are based either on subjective weights, discriminant analysis, logit and probit estimations, principal components.

Residual value model or option pricing approach. Modigliani and Miller Theorem implies that the value of a firm must add up to the value of its loan instruments plus equities. When equities are negative, you default on loans. Problems with predictability (first-passage time). Modelling jumps and unpredictable default time.

Reduced form models or models based upon the no-arbitrage hypothesis. Problems with identifying equivalence classes when estimating the term structure of default risk. Need for clearly identifiable market benchamarks. Use of Markov chains in modelling transition (migration) from one class of risk to the other. Conditional Markov chains. Conditional on what? Problems with estimation of systematic risk.

Market risk and credit risk are strictly related. Market risk measures, in some sense, the systematic risk component. Market fluctuations have a causal impact upon credit riskiness. This is most obvious with derivative instruments: a swap can suddenly turn a lender into a borrower, thereby generating credit risk which did not previously exist.

1.1 Decomposing credit risk

Denote with L the potential loss. This can be broken down into three parts

i. binary event of a missed payment, $D \in \{0, 1\}$;

ii. total exposure $X \in \Re_+$, upon which D operates;

iii. the recovery rate $R \in [0, 1]$.

$$L = D \times X \times (1 - R) \tag{1.1}$$

Notice that X is not fixed once and for all. It is so only for straight bonds. Bank loans offer an interval of values up to the maximum available line of credit. Derivative instruments are characterized by a value of X that is highly variable, can move from negative to positive and back and is dependent upon market's movements.

How about D and R? In principle you go out and measure them from historical data. But how?

1.2 Empirical aspects of credit risk

It is important to distinguish between the book value and the economic value of a loan. The relevant concept for our purposes is the economic value, and we shall refer to this concept simply as the value of the loan.

Example 1 Consider a loan with a book value of \$100, residual life of 5 years, paying a fixed annual interest rate of 6%. Its (economic) value is

$$Y = \sum_{t=1}^{5} \frac{6}{(1+r_t)^t} + \frac{100}{(1+r_5)^5},$$

where r_t is a discount rate that depends on the time when payments are made and the riskiness of the loan. For example, if r_1^* denotes the interest rate on a one-year zero-coupon Treasury bond, π_1 denotes the default probability over a one-year horizon and R_1 denotes the amount recovered in case of default, then we have the following arbitrage condition

$$(1-\pi_1)(1+r_1) + \pi R_1 = 1 + r_1^*,$$

from which we get

$$r_1 = \frac{1 + r_1^* - \pi_1 R_1}{1 - \pi_1} - 1.$$

Two loans with the same book value may have very different economic value depending on their residual life, the interest rate that they pay in each period, and the set of discount rates adopted. □

By credit risk we mean the uncertainty in the value of the portfolio at a given risk horizon caused by the possibility of obligor credit quality changes (upgrades, downgrades and default). It is important to stress the difference between expected (or mean) loss

$$\text{Expected loss} = E(\text{expected size of loss} \mid \text{loss}) \times \Pr\{\text{loss}\}$$

and risk, that is, the probability distribution (or "volatility") of the value of a credit portfolio. In the case of credit risk, the distinction is important because:

i. credit-related losses are uncertain,

ii. their distribution is heavily skewed, which makes it inadequate to describe risk through a few summaries such as the mean and the standard deviation.

The mean is a measure of center that works well for symmetric distributions but is very sensitive to outliers. For asymmetric or heavy-tail distributions, other measures (such as the median or the mode) may be more appropriate. The standard deviation is a symmetric measure of spread and is even more sensitive to outliers than the mean.

For heavy-tail distributions, the probability of extreme events is higher than for thin-tail ones. Because of this, much more attention needs to be devoted to understanding tail behavior relative to behavior in the middle of the distribution.

Whether credit events should be treated as discrete or continuous random variables is an issue that depends both on the specific situation being modelled and considerations of tractability, analytical simplicity, etc. We do not want to take a particular position on this issue. More pragmatically, we shall adopt the approach that we feel best suits the problem at hand.

Other important issues that arise in practice are:

- Precise definition of default.
- Specification of the time when the loss due to default is to be computed.
- Choice of risk horizon: the convention is one year, but shorter or longer periods may also be considered.
- How to compute exposures and recovery rates.
- How to incorporate information about the characteristics of the obligor, such as balance-sheet indicators, calendar time and macroeconomic variables.
- How to allow for parameter uncertainty. This issue arises because key parameters are often not known and must be estimated from sample data or chosen in some other way.
- How to allow for model errors. This issue arises because models are at best an approximation to reality.

Representing the volatility of portfolio value is typically done in steps. The first step consists in obtaining the distribution of values for a single loan (e.g. obtaining the range of values that a bond of a given quality can take at the end of a period and the associated probabilities).

The second step consists in obtaining the distribution of values for a portfolio with two or more loans. It is worth noticing right now that, in a portfolio of N

loans with S states of credit quality for each loan, we need to attach probabilities to S^N states. Even for moderate N and S, this is a large number, thus raising serious computational problems.

Computational problems are likely to be particularly serious in case of banks, characterized by a high number of small loans. In this case, one may try to work with "synthetic obligors", that is, very fine and sufficiently homogeneous classes which may be treated as individual obligors.

1.3 International Regulations

1988 Basle Accord and successive improvements, addendum and revisions. Last one in 1999.

See BIS site for information, details and technical documents.

Basic tension is between a in-house, model based evaluation of CR and a standardized and uniformed evaluation from outside.

The latter is, obviously, too risk adverse and punitive, i.e. it exagerates the capital requirements that banks are expected to meet.

Moral hazard problems are gigantic.

Classical tradeoff between economic efficiency and stability of the system.

Regulator emphasizes stability and has a tendency to "eliminate" risk by imposign very stringent capital requirements.

Due to the existence of various insurance mechanisms and national and international lenders of last resort, banks have an incentive to take up more risk than it is socially optimal.

Systemic risk: myth or real danger? In spite of the many claims in the popular press and of the attention the issue has received by politicians, regulators and international organizations, on pure scientific grounds there are few if any reasons to believe that the danger of "systemic collapse" is real. As such, there are very few reasons to support the idea that regulators should impose strict and direct capital requirements and portfolio compositions rules upon banks, as opposed to just enforcing transparency and full release of information relative to the composition and riskiness of their investment portfolio.

2 Pricing

2.1 Arbitrage: basic notions

Three instruments, X^1, X^2, X^3.

First instrument has high and low return. Second has fixed return R and

$$H > R > L. \tag{2.2}$$

Third has return R^h and R^l respectively. We are interested in the restrictions that the principle of no arbitrage, see below, imposes on such returns and prices of the three assets.

Current prices are p^1, p^2 and p^3.

	X^1	X^2	X^3
h	H	R	R^h
l	L	R	R^l
	p^1	p^2	p^3

(2.3)

In this context, a portfolio is a triple (x^1, x^2, x^3) of real numbers summing up either to zero (self-financing) or to one (positive wealth)

Assume that for a certain triple $(\hat{x}^1, \hat{x}^2, \hat{x}^3)$

$$\hat{x}^1 H + \hat{x}^2 R + \hat{x}^3 R^h \geq 0 \tag{2.4}$$

$$\hat{x}^1 L + \hat{x}^2 R + \hat{x}^3 R^l \geq 0. \tag{2.5}$$

Is this consistent with

$$p^1 \hat{x}^1 + p^2 \hat{x}^2 + p_3 \hat{x}^3 < 0? \tag{2.6}$$

It is not, because if it were one could

i. purchase $\hat{x}$,

ii. realize a strictly positive profit today,

iii. have nothing to pay tomorrow.

This is the principle of no arbitrage: any portofolio with non-negative payoff tomorrow in all states of the world, must have a non-negative price today.

2.2 Implications of no-arbitrage

One can prove that, whenever arbitrage is absent then "probabilities" can be assigned to future states of the world. Note we did not talk about the probability of high and low until now. If no-arbitrage (NA) is satisfied, then there exists numbers y^h and y^l such that

$$y^h \geq 0, y^l \geq 0, \tag{2.7}$$

and

$$\begin{aligned} y^h H + y^l L &= p^1, \\ (y^h + y^l) R &= p^2, \\ y^h R^h + y^l R^l &= p^3. \end{aligned} \tag{2.8}$$

Let $Y \equiv y^h + y^l$. Equations 2.8, imply that $Y > 0$, hence we can define

$$q = \frac{y^h}{Y}, 1 - q = \frac{y^l}{Y}. \tag{2.9}$$

Hence q can be treated as a probability or, in the jargon, "pseudo probability" or risk free probability.

Let us manipulate 2.8 a little bit

$$Y = \frac{p^2}{R}. \tag{2.10}$$

and

$$qH + (1-q)L = \frac{p^1 R}{p^2}. \tag{2.11}$$

Hence

$$q = \frac{p^1 R - p^2 B}{p^2 (H - L)}. \tag{2.12}$$

Hence both Y and the pseudo probability q are contained or implied by equilibrium prices and returns.

Furthermore, the third equation 2.8 also implies that

$$p^3 = (qR^h + (1-q)R^l)\frac{p^2}{R}. \tag{2.13}$$

Under NA, redundant securities can be priced using their returns and risk neutral probabilities by simple application of the expected value formula. This leads to the notion of "redundant securities". Basic intuition is simply a counting of equations and unknown. i.e. of securities and states of nature. If there are more (independent: i.e. without collinear payoff vectors) securities than states of nature, then the additional securities are "redundant" and can be priced, because of NA, by means of the prices of the other securities.

In our example we have two states of nature, and three independent securities. Hence, one of them is redundant. In particular, it is clear that there exists weights (x^1 and x^2) for the first and second security such that

$$\begin{aligned} x^1 H + x^2 R &= R^h \\ x^1 L + x^2 R &= R^l. \end{aligned} \tag{2.14}$$

By inverting a matrix (for this, independence of returns is requested)

$$\begin{aligned} x^{*1} &= \frac{R^h - R^l}{H - L} \\ x^{*2} &= \frac{HR^l - LR^h}{R(H-L)}. \end{aligned} \tag{2.15}$$

This solution gives another way of thinking about arbitrage. The portfolio (x^{*1}, x^{*2}) can be thought of as a security. By definition its price must be

$$p^1 x^{*1} + p^2 x^{*2};$$

and it must be priced as X^3, because it gives the same identical payoff in all states of nature, so

$$p^3 = p^1 x^{*1} + p^2 x^{*2}; \tag{2.16}$$

2.3 No Arbitrage in general

Given securities X^i, with $i = 1, \dots n$, and states $s = 1, \dots m$. Denote returns with $R^i(s)$ and prices with p^i. Let the first security be the risk free security

$$R^1(s) = R \text{ for all } s.$$

No arbitrage here means that a portoflio $(x^i), i = 1, \dots, n$ satisfying

$$\begin{aligned} \sum_{i=1}^{n} R^i(s) x^i &\geq 0 \\ \sum_{i=1}^{n} p^i x^i &< 0 \end{aligned} \tag{2.17}$$

is impossible.

Then we conclude that there exists a vector of pseudo or risk free probabilities $(q^s)_{s=1,\dots,m}$ such that

$$\frac{1}{R}\sum_{s=1}^{m} q^s R^i(s) = \frac{p^i}{p^1}, \tag{2.18}$$

for all $i = 1, \dots, n$

Assume now that

$$n = m + \overline{n}, \qquad \overline{n} > 0.$$

Then

- the pseudo probabilities are uniquely determined by the first m equations in 2.18;
- the prices of the $\overline{n}$ redundant securities are determined by those of the first m securities..

Theorem *Let the real numbers $D^i(s)$, with $i = 1, 2, \dots, n$ and $s = 1, 2, \dots, m$, and the vector $p = (p^1, \dots, p^i, \dots p^n)$ be given. Then, for all $x = (x^1, x^2, \dots, x^n)$*

$$\sum_{i=1}^{n} D^i(s)x^i \geq 0 \tag{2.19}$$

for all s implies that

$$\sum_{i=1}^{n} p^i x^i \geq 0 \tag{2.20}$$

if and only if there exists a vector $y = (y^1, y^2, \dots, y^m)$ such that $y^s \geq 0$ for all s, and

$$\sum_{s=1}^{m} y^s D^i(s) = p^i \tag{2.21}$$

for all i.

2.3.1 General Pricing Formula

Given the return and states of nature structure

s	$p(t+1, T, (s^t, s))$	$p(t+1, T', (s^t, s))$	$r(s^t)$
...	...	...	...
s'	$p(t+1, T, (s^t, s'))$	$p(t+1, T', (s^t, s'))$	$r(s^t)$
	$p(t, T, s^t)$	$p(t, T', s^t)$	1

(2.22)

The price $p(t, T, s^t)$, is

$$\frac{1}{r(s^t)}\sum_{s\in S} q^s p(t+1, T, (s^t, s)) = p(t, T, s^t). \tag{2.23}$$

which is just a fancy way of writing the net present value formula.

2.4 Pricing with default

Begin by introducing the, idiosyncratic, events f = default, and n = no-default.

$$\omega \equiv (s, x), x \in \{n, f\} \tag{2.24}$$

Assume f is an absorbing state. From n the Markov chain can take you to either n, with probability $1-q$, or f, with probability q.

Security pays one if state is n and δ, with $0 < \delta < 1$ if state is f. The price is $v(t, T, \omega))$. In the simplest case there is no uncertainty in the aggregate state of nature, so the interest rate is fixed and the only random variable is the idiosyncratic state of the counterpart (borrower). Compute $v(t, T, x)$ for all t and x.

Use the recursive method on the present value formula. At expiration

$$v(T, T, f) = \delta, v(T, T, n) = 1. \tag{2.25}$$

Period before the last:

n	1	r
f	δ	r
	$v(T-1, T, n)$	1

(2.26)

$$v(T-1, T, n) = \frac{1}{R}[q\delta + (1-q)] \tag{2.27}$$

$$v(T-1, T, f) = \frac{\delta}{R}. \tag{2.28}$$

For a generic period t we have

n	$v(t+1, T, n)$	R
f	$v(t+1, T, f)$	R
	$v(t, T, n)$	1

(2.29)

Hence

$$v(t, T, n) = \frac{1}{R^{T-t}}\left([\delta(q + q(1-q) + \dots (1-q)^{T-t-1})] + (1-q)^{T-t}\right) \tag{2.30}$$

and

$$v(t, T, f) = \frac{\delta}{R^{T-t}} \tag{2.31}$$

Recall that

$$q + q(1-q) + \dots (1-q)^{T-t-1}$$

is the probability of default between period t and T, while $(1-q)^{T-t}$ is the probability that default never takes place in the periods from t to T.

Also in this simple case the NA condition can be applied, to yield

$$v(t,T,n) = \frac{1}{R}[qv(t+1,T,f) + (1-q)v(t+1,T,n)] \tag{2.32}$$

and

$$v(t,T,f) = \frac{1}{R}v(t+1,T,f), \tag{2.33}$$

for all t.

Let us now move on to the case in which the state of nature has an aggregate component, i.e. $\omega = (s,x) \in \{h,l\} \times \{n,f\}$. We write the price as $v(t,T,(s,x))$. Apply the backward induction present value method once again

$$v(T,T,(s,n)) = 1, v(T,T,(s,f)) = \delta, \tag{2.34}$$

for all s. For the previous period

n	1	$r(s)$
f	δ	$r(s)$
	$v(T-1,T,(s,n))$	1

(2.35)

$$v(T-1,T,(s,n)) = \frac{1}{r(s)}[q\delta + (1-q)], \tag{2.36}$$

$$v(T-1,T,(s,f)) = \frac{\delta}{r(s)}. \tag{2.37}$$

For a generic period t

(s',n)	$v(t+1,T,(s',n))$	$r(s)$
(s',f)	$v(t+1,T,(s',f))$	$r(s)$
	$v(t,T,(s,n))$	1

(2.38)

Hence it is natural to conjecture that the pricing formula satisfies

$$v(t,T,(s_t,n)) = \tag{2.39}$$

$$\frac{1}{r(s_t)} \sum_{\{s_{t+1},\ldots,s_{T-1}\}} \frac{q(s_{t+1})\ldots q(s_{T-1})}{r(s_{t+1})\ldots r(s_{T-1})} \left([\delta(q + q(1-q) + \ldots (1-q)^{T-t-1}] + (1-q)^{T-t}\right)$$

and

$$v(t,T,(s_t,f)) = \frac{1}{r(s_t)} \sum_{\{s_{t+1},\ldots,s_{T-1}\}} \frac{q(s_{t+1})\ldots q(s_{T-1})}{r(s_{t+1})\ldots r(s_{T-1})}\delta \tag{2.40}$$

3 Statistical Notions

3.1 Representing standalone risk

We begin with the problem of representing the volatility of a single loan in a credit portfolio, that is, the variability of its end-of-period value. Mathematically, volatility may be represented by a real valued random variable (r.v.) Y. Because the loss is the random quantity

$$L = Y_0 - Y,$$

where Y_0 denotes the nominal value of the loan, risk may be fully assessed if we know the probability distribution of Y.

The simplest possible case is when there is a fixed default probability π, a fixed exposure ϵ and a fixed recovery value R which, for simplicity, we set equal to zero. In this case we may work directly with the loss L, as

$$L = \begin{cases} \epsilon, & \text{with probability } \pi, \\ 0, & \text{with probability } 1 - \pi. \end{cases}$$

Because of the Bernoulli nature of the problem, the risk is completely characterized by the mean and variance of the loss

$$E(L) = \epsilon\pi, \qquad \mathrm{Var}(L) = E(L^2) - [E(L)]^2 = \epsilon^2\pi(1 - \pi).$$

Notice that the mean and the variance of L are related through

$$\mathrm{Var}(L) = \frac{1 - \pi}{\pi}\,[E(L)]^2.$$

If a sufficiently wide range of end-of-period values are possible, then it may be sensible to treat the value Y (and therefore the loss L) as a continuous r.v.

3.2 Predicting default

Formally, the problem consists of classifying a loan into one of two mutually exclusive groups (default and no-default) based on a K-vector of observed characteristics x. Qualifications should be added to this prediction: time horizon and extent of default.

Prediction has been based on a variety of models such as logit and probit models, discriminant analysis, and semi-nonparametric techniques such as neural networks. None of the above methods emerges as a clear winner.

3.2.1 Logit and probit

Let $X = (X_1, \ldots, X_K)$ be an observable random K-vector that represents all we know about the loan (characteristics of the obligor, the country/region and sector of activity, the calendar time, macroeconomic variables, etc.) Logit and probit are alternative ways of modeling the conditional default probability

$$\pi(x) = \Pr\{\text{default} \mid X = x\},$$

where $x = (x_1, \ldots, x_K)$ is one of the possible values of X.

In the *logit* approach, we introduce a linear combination or "index"

$$\gamma + \sum_{k=1}^{K} \delta_k X_k = \gamma + x^\top \delta,$$

where γ is a scalar parameter and δ is a K-vector of parameters, and model the conditional default probability as

$$\pi(x) = \frac{\exp(\gamma + x^\top \delta)}{1 + \exp(\gamma + x^\top \delta)} \tag{3.41}$$

or, equivalently in terms of log odds-ratio,

$$\ln \frac{\pi(x)}{1 - \pi(x)} = \gamma + x^\top \delta.$$

In the *probit* approach, the conditional default probability is instead modelled as

$$\pi(x) = \Phi(\gamma' + x^\top \delta'),$$

where $\Phi(u)$ denotes the standard normal d.f. Notice that $F(u) = \exp u/(1 + \exp u)$ is the unit logistic d.f.

There exists a simple relationship between the parameters $\theta = (\gamma, \delta)$ of the logit model and the paramaters $\theta' = (\gamma', \delta')$ of the probit model, for $\theta' \approx 1.81 \times \theta$. This is because: (i) the standard normal and the unit logistic have a very similar shape and differ only in the tails, which are slightly heavier for the logistic, and (ii) the standard normal has unit variance, whereas the logistic distribution has variance equal to $3.14^2/3$.

Notice that, in both cases, the linear index is continuously mapped into the interval $[0, 1]$, which is the range of values for probabilities. Thus, more generally, we may consider

$$\pi(x) = F(\gamma + x^\top \delta),$$

where F is any continuous function from $[0,1]$ into $\Re$.

The logit and probit models may be interpreted as arising from the following simple threshold model. Let D be an indicator of default, equal to 0 if default occurs and equal to 1 otherwise. Let the continuous r.v. Y^* denote the unobservable end-of-period value of the obligor and let F be the d.f. of Y^*. We assume that default occurs whenever Y^* falls below some cut-off value $C(x)$ which may depend on the characteristics of the obligor, that is, $D = 0$ whenever $Y^* < C(x)$. The conditional default probability is therefore

$$\pi(x) = \Pr\{D = 0 \mid x\} = \Pr\{Y^* < C(x)\} = F(C(x)),$$

where we used the fact that F is continuous. Let $C(x) = \gamma + x^\top \delta$. Then we obtain the logit model if F is logistic and the probit model if F is normal.

There is another possible interpretation of the threshold model. Suppose that the latent value may be represented as

$$Y^* = \mu(x) + \sigma U,$$

where $\mu(x)$ depends on the characteristics of the obligor and U is a r.v. with d.f. F. The conditional default probability is then equal to

$$\pi(x) = \Pr\{Y^* < C \mid x\} = \Pr\{\mu(x) + \sigma U < C\} = F\left(\frac{C - \mu(x)}{\sigma}\right).$$

Let $\mu(x) = C - \sigma(\gamma + x^\top \delta)$. As before, we obtain the logit model if F is the unit logistic d.f. and the probit model if F is the standard normal d.f.

The second interpretation of the threshold model implies

$$\frac{\partial \pi}{\partial C} = \frac{1}{\sigma} f\left(\frac{C - \mu(x)}{\sigma}\right) = -\frac{\partial \pi}{\partial \mu}$$

and

$$\frac{\partial \pi}{\partial \sigma} = -\frac{1}{\sigma} f\left(\frac{C - \mu(x)}{\sigma}\right) \cdot \frac{C - \mu}{\sigma},$$

where f denotes the density of U. Notice that, while $\partial \pi / \partial C > 0$ and $\partial \pi / \partial \mu < 0$, the sign of $\partial \pi / \partial \sigma$ is positive or negative depending on whether $C < \mu$ or $C > \mu$. Since $C < \mu$ in general, we should expect $\partial \pi / \partial \sigma > 0$.

Given knowledge of the parameters (γ, δ) (they may be estimated from the data using various methods, such as nonlinear least squares or maximum likelihood), one may compute the default probability of a new loan with characteristics x. The conditional default probability $\pi(x)$ may then be used to assign a loan to one of $S \geq 2$ credit quality classes.

In practice, this is done by partitioning the interval $[0, 1]$ into a set of bins defined by cut-off points

$$0 = C_0 < C_1, \ldots, C_{S_1} < C_S = 1.$$

A loan with characteristics x is assigned credit grade d if $\pi(x)$ falls in the dth bin, that is, $C_{d-1} \leq \pi(x) < C_d$. In the empirical implementation of the method, the choice of the number and position of the cut-offs is essential and must be decided by considering the trade-off between precision, computational time and the probability of misclassification.

The important modeling issues are:

i. the choice of F;

ii. given the choice of F, how to model the dependence of π on x (this includes the issue of how to model time-dependence);

iii. modeling unobserved heterogeneity, that is, differences in default probabilities that cannot be explained by x.

One advantage of the threshold model is that it gives an indication on how to choose F for, in this model, F is simply the d.f. of the end-of-period value of the obligor. Notice, however, that this distribution need not be logistic or normal.

Turning to the problem of how to model the dependence of π on x, in addition to the use of a linear index, we may consider more flexible specifications, such as polynomial functions in x.

Finally, unobserved heterogenity may be modeled by introducing an unobservable r.v. ν (whose distribution may depend on a vector of parameters ψ) such that

$$\pi(x \mid \nu) = F(\nu + x^\top \delta).$$

If $g(v; \psi)$ denotes the density function of ν, then

$$\pi(x) = \int F(v + x^\top \delta)\, g(v; \psi)\, dv.$$

Notice that, even if F is logistic or normal, the resulting model for $\pi(x)$ is not logit or probit in general.

3.2.2 Discriminant analysis

Suppose that there are two alternative credit states: default and no-default. Let the density of the random K-vector X be $f_0(x)$ in the first class (default) and $f_1(x)$ in the second (no-default). Also let π be the unconditional probability of default, estimated by the fraction of loans belonging to the first class.

We seek a partition of $\Re^K$ into two regions, R_0 and R_1, such that, if X falls in R_h, then we classify the loan into class $h = 0, 1$. For any proposed partition, misclassification is possible as "default" loans may be classified as "no-default" and viceversa. These two events occur with probability $\pi \Pr\{X \in R_1\}$ and $(1-\pi) \Pr\{X \in R_0\}$ respectively. An optimal partition is one that minimizes the expected total cost of misclassification.

If C_h denotes the cost of incorrectly classifying a loan into group h, then the expected total cost of misclassification is

$$\begin{aligned} C &= C_0\pi \Pr\{X \in R_1\} + C_1(1-\pi)\Pr\{X \in R_0\} \\ &= C_0\pi(1 - \Pr\{X \in R_0\}) + C_1(1-\pi)\Pr\{X \in R_0\} \\ &= C_0\pi + \int_{R_0} [C_1(1-\pi)f_1(x) - C_0\pi f_0(x)]\, dx. \end{aligned}$$

This is minimized if R_0 consists of all points such that

$$C_1(1-\pi)f_1(x) - C_0\pi f_0(x) < 0,$$

or equivalently

$$\frac{f_0(x)}{f_1(x)} > \frac{C_1}{C_0}\frac{1-\pi}{\pi}. \tag{3.42}$$

Normal discriminant analysis corresponds to the case when X is distributed as $\mathcal{N}(\mu_h, \Sigma_h)$ with density

$$f_h(x) = c \times |\Sigma_h|^{-1/2} \exp[-\frac{1}{2}(x-\mu_h)^\top \Sigma_h^{-1}(x-\mu_h)], \qquad h = 0, 1.$$

In the special case when $\Sigma_0 = \Sigma_1 = \Sigma$, then

$$\begin{aligned} \frac{f_0(x)}{f_1(x)} &= \exp[-\frac{1}{2}(x-\mu_0)^\top \Sigma^{-1}(x-\mu_0) + \frac{1}{2}(x-\mu_1)^\top \Sigma^{-1}(x-\mu_1)] \\ &= \exp[(\mu_0-\mu_1)^\top \Sigma^{-1} x - \frac{1}{2}(\mu_0-\mu_1)^\top \Sigma^{-1}(\mu_0+\mu_1)]. \end{aligned}$$

Hence, after taking logs, condition (3.42) becomes

$$(\mu_0-\mu_1)^\top \Sigma^{-1} x > \ln\frac{C_1(1-\pi)}{C_0\pi} + \frac{1}{2}(\mu_0-\mu_1)^\top \Sigma^{-1}(\mu_0+\mu_1).$$

Defining

$$\delta = \Sigma^{-1}(\mu_0 - \mu_1), \tag{3.43}$$

the optimal choice of R_0 (the default region) is defined by the condition

$$\delta^\top x > \ln \frac{C_1(1-\pi)}{C_0\pi} + \delta^\top \bar{\mu},$$

where $\bar{\mu} = (\mu_0 + \mu_1)/2$. The linear combination $\delta^\top x$ is known as *linear discriminant function* or *Z-score.*

In the special case when $C_1(1-\pi) = C_0\pi$, the cut-off for the Z-score is equal to $\delta^\top \mu$. Given δ and $\bar{\mu}$, we then assign a loan with characteristics x to the default class whenever $\delta^\top(x - \bar{\mu})$ is positive. The other crucial assumptions behind this result are:

i. Both $f_0(x)$ and $f_1(x)$ are multivariate normal. If they are not, then the basic model remain valid but the expression for the likelihood ratio $f_0(x)/f_1(x)$ is more complicated.

ii. The covariance matrices are equal. If the model is normal but this assumption is violated, then the discriminant function becomes quadratic in x, that is, of the form $x^\top \Delta x + \delta^\top x$.

3.2.3 Relationship between discriminant analysis and logit

Approaching the problem from a Bayesian viewpoint gives a relationship between linear discriminat analysis and the logit model. By Bayes Theorem, the conditional (or posterior) probability of "default" given $X = x$ and the prior probability π is

$$\pi(x) = \Pr\{\text{default} \mid X = x\} = \frac{f_0(x)\pi}{f_0(x)\pi + f_1(x)(1-\pi)}.$$

If $f_h(x)$ is the multivariate normal density with mean μ_h and variance Σ, then the posterior odds-ratio is

$$\frac{\pi(x)}{1-\pi(x)} = \frac{f_0(x)}{f_1(x)} \frac{\pi}{1-\pi} = \exp(\gamma + x^\top \delta),$$

where

$$\gamma = \ln \frac{\pi}{1-\pi} - \frac{1}{2}(\mu_0 - \mu_1)^\top \Sigma^{-1}(\mu_0 + \mu_1)$$

and δ is given by (3.43). This is just the logit specification (3.41).

3.2.4 Semi-nonparametric models

This kind of models tries to avoid the parametric assumptions made by (linear) logit and probit models while placing some structure on default probabilities in order to avoid the "curse of dimensionality" problem that makes fully nonparametric methods inpractical unless $K \leq 3$ or 4.

One example is the *additive logit model*

$$\ln \frac{\pi(x)}{1-\pi(x)} = \gamma + \sum_{k=1}^{K} h_k(x_k),$$

where $h_1, \ldots, h_K$ is a set of arbitrary smooth univariate functions, one for each component of X.

Another example is the *feedforward single-layer neural network*

$$\ln \frac{\pi(x)}{1-\pi(x)} = \beta_0 + \sum_{j=1}^{q} G(x^\top \delta_j)\, \beta_j, \tag{3.44}$$

where G is a known univariate function. We obtain the logit model when $q = 1$ and G is the identity function.

3.3 A general framework

We now consider generalizations of the model in Section 3.1 along three directions:

i. we allow for more than two credit states, which enables us to incorporate information about credit quality changes;

ii. for each credit state, we allow for variability of the end-of-period value of each loan in the portfolio;

iii. we allow for risk horizons longer than one year.

3.3.1 Modeling the risk of a single loan

We begin with the case when the time unit is the year and the risk horizon is one year, and then consider the case when risk is measured with reference to a longer horizon of $h \geq 1$ periods.

Let the r.v. Y_{t+1} represent the variability of the end-of-period value of a given loan in a portfolio. It helps things considerably if we know something about the loan, in addition to its amount. Here we assume that we know its current credit quality state (the extension to the case where we also observe a vector X of individual and macroeconomic indicators is straightforward). We also assume that there is a finite number $S \geq 2$ of mutually exclusive and totally exhaustive credit states, one of which is the state of default taken to be absorbing and labelled by 0.

An important practical question, which we do not discuss at this stage, is how many credit states there are and how they have been defined. For concreteness, we assume that they are based on some externally given rating system, such as Moody's, S&P or KMV.

Let D_t be an indicator of credit state at time t, with default corresponding to $D = 0$. We want to use the information that $D_t = d$ for the given loan to determine the distribution of its forward value.

The variability of the end-of-period value of the loan given credit state $D_t = d$ at time t is represented by the "predictive density"

$$f_{t+1|t}(y \mid d) = \sum_{d'} f(y \mid D_{t+1} = d', D_t = d)\, \pi(d' \mid d), \tag{3.45}$$

where $f(y \mid D_{t+1} = d', D_t = d)$ is the conditional probability density of Y_{t+1} given credit states $D_{t+1} = d'$ and $D_t = d$ at times $t+1$ and t respectively, and

$$\pi(d' \mid d) = \Pr\{D_{t+1} = d' \mid D_t = d\}$$

is the transition probability (or "migration likelihood") from state d to state d' over a one-year period. We assume, for simplicity, that $\pi(d' \mid d)$ is time invariant. If we assume that default is an absorbing state, then $\pi(0 \mid 0) = 1$ and $\pi(d' \mid 0) = 0$ for all other d'.

Assuming that

$$f(y \mid D_{t+1} = d', D_t = d) = f(y \mid D_{t+1} = d'), \tag{3.46}$$

the density (3.45) becomes

$$f_{t+1|t}(y \mid d) = \sum_{d'} f(y \mid d')\, \pi(d' \mid d),$$

where $f(y \mid d') = f(y \mid D_{t+1} = d')$. Assumption (3.46) is another stationarity assumption. It also requires the distribution of Y to depend only on the current credit state. Notice that the predictive density is simply a weighted average of the conditional densities $f(y \mid d')$, with weights given by the migration likelihoods $\pi(p' \mid p)$.

Now let the r.v. Y_{t+h} represent the variability of the value of the loan h period ahead, with $h \geq 1$. By the same argument used before, the conditional density of Y_{t+h} given $D_t = d$ is

$$f_{t+h|t}(y \mid d) = \sum_{d'} f(y \mid d')\, \pi^{(h)}(d' \mid d),$$

where $f(y \mid d') = f(y \mid D_{t+h} = d')$ and

$$\pi^{(h)}(d' \mid d) = \Pr\{D_{t+h} = d' \mid D_t = d\}$$

is the transition probability from state d to state d' over a h-year period.

3.3.2 Modeling "migrations"

In this section we look at ways of using the knowledge of a vector of observed characteristics of the loan to predict its "migration likelihood". To simplify notation, we absorbe the knowledge of the current credit state into the vector X_t and drop the time subscripts from X_t and D_{t+1}.

A simple possibility is the multinomial logit (MNL) specification

$$\pi(0 \mid x) = \Pr\{D = 0 \mid X = x\} = \frac{1}{1 + \sum_{d=1}^{S-1} \exp(\alpha_d + x^\top \beta_d)}$$

and

$$\pi(d \mid x) = \Pr\{D = d \mid X = x\} = \frac{\exp(\alpha_d + x^\top \beta_d)}{1 + \sum_{d=1}^{S-1} \exp(\alpha_d + x^\top \beta_d)}, \qquad d = 1, \ldots, S-1,$$

where the parameters α_d and β_d depend on both the future state d and the current state. More compactly, the MNL may be represented in terms of log-odds as

$$\ln \frac{\pi(d \mid x)}{\pi(0 \mid x)} = \alpha_d + x^\top \beta_d, \qquad d = 1, \ldots, S-1,$$

where we take state 0 (default) as the baseline or reference state.

Although attractive for its simplicity, the MNL specification places strong restrictions on credit state probabilities, for it implies that the log odds

$$\ln \frac{\pi(d' \mid x)}{\pi(d \mid x)} = (\alpha_d - \alpha'_d) + x^\top (\beta_d - \beta'_d),$$

depend only on two states being compared and not on the other credit states. This property, known as *independence of irrelevant alternatives* (IIA), may lead to unreasonable conclusions in certain cases. An alternative to MNL, discussed in Section 4.3.1 and implemented in J.P. Morgan (1997), is the ordered probit model.

One-period transition probabilities may be used to generate h-period ahead transition probabilities by expoiting the basic recurrence relationship

$$\pi_{t+1} = \Pi^\top \pi_t, \tag{3.47}$$

where π_t is the S-vector of credit state probabilities at time t and $\Pi = [\pi(d' \mid d)]$ is the $S \times S$ matrix of transition probabilities over a one-year period. If transition probabilities are time invariant, then iterating (3.47) we get

$$\pi_{t+h} = \Pi^{(h)^\top} \pi_t, \qquad h = 1, 2, \ldots,$$

where $\Pi^{(h)} = \Pi^h$ is the matrix of transition probabilities over a h-year period.

If transition probabilities are not time-invariant, that is, the elements of the matrix Π depend on time, then

$$\Pi_t^{(h)} = \prod_{j=1}^{h} \Pi_{t+j}.$$

3.4 Portfolio credit risk

We begin with a simple model with only two credit states (default and no-default) and a one-year risk horizon. The bank portfolio consists of N loans (indexed by $j = 1, \ldots, N$), each with a fixed default probability π_j, a fixed exposure ϵ_j and a fixed recovery value R_j which, for simplicity, we set equal to zero. The number of defaults is random and can vary anywhere between 0 and N. The portfolio loss is the sum of all random individual losses, $L = \sum_{j=1}^{N} L_j$, where the distribution of L_j was derived in Section 3.1.

The expected portfolio loss is easy to compute, being simply the sum of the individual expected losses

$$E(L) = \sum_{j=1}^{N} E(L_j) = \sum_{j=1}^{N} \epsilon_j \pi_j.$$

It is intuitively clear, on the other hand, that the variance of the total loss depends on the standalone risks plus the correlations between the standalone risks. In the special case when the standalone risks are independent we have

$$\mathrm{Var}^*(L) = \sum_{j=1}^{N} \mathrm{Var}(L_j) = \sum_{j=1}^{N} \epsilon_j^2 \pi_j (1 - \pi_j).$$

In general, however,

$$\mathrm{Var}(L) = \mathrm{Var}^*(L) + 2 \sum_{j=1}^{N} \sum_{i>j} \mathrm{Cov}(L_i, L_j),$$

where

$$\mathrm{Cov}(L_i, L_j) = \epsilon_i \epsilon_j (\pi_{ij} - \pi_i \pi_j)$$

and π_{ij} denotes the probability that i and j both default. If defaults are positively correlated, e.g. because they are influenced by common economic conditions, then $\pi_{ij} \geq \pi_i \pi_j$ and we would expect $\mathrm{Var}(L) > \mathrm{Var}^*(L)$. The opposite may be true if defaults are negatively correlated, that is, $\pi_{ij}^* < \pi_i \pi_j$.

Notice that, except in the case of independence, knowledge of π_{ij}^* is essential in order to compute the variance of the portfolio. With N loans, there are $N(N+1)/2$

joint probabilities and the important practical problems arises of how to estimate such a large number of probabilities from the available data.

For risk calculations we also need the probability distribution of L. This is the distribution of a sum of potentially correlated Bernoulli r.v.s. with mean $\mu_j = \epsilon_j \pi$, variance $\sigma_j^2 = \epsilon_j^2 \pi_j (1-\pi_j)$ and covariance $\sigma_{ij} = \epsilon_i \epsilon_j (\pi_{ij}^* - \pi_i \pi_j)$. The exact distribution of L is easily computed for small N.

Example 2 Suppose that $N = 2$, let π_{00} denote the probability that both loans default, π_{01} the probability that only the first defaults, π_{10} the probability that only the second defaults, and π_{11} the probability that none of them defaults. The distribution of the portfolio loss is then

$$L = \begin{cases} \epsilon_1 + \epsilon_2, & \text{with probability } \pi_{00}, \\ \epsilon_1, & \text{with probability } \pi_{01}, \\ \epsilon_2, & \text{with probability } \pi_{10}, \\ 0, & \text{with probability } \pi_{11}. \end{cases}$$

Because the marginal default probabilities of the two loans are $\pi_1 = \pi_{00} + \pi_{01}$ and $\pi_2 = \pi_{00} + \pi_{10}$, it is easily verified that

$$E(L) = \epsilon_1 \pi + \epsilon_2 \pi_2$$

and

$$\text{Var}(L) = \epsilon_1^2 \pi_1 (1 - \pi_1) + \epsilon_2^2 \pi_2 (1 - \pi_2) + 2\epsilon_1 \epsilon_2 (\pi_{12}^* - \pi_1 \pi_2),$$

where $\pi_{12}^* = \pi_{00}$. □

When N is large, one may either rely on simulations or work with suitable approximations. In the special case when $L_1, \ldots, L_N$ are independent, we have the following Central Limit Theorem (CLT):

Theorem 3.1 (Lindeberg-Feller) *Let $\{L_j\}$ be a sequence of independent r.v.s with mean μ_j, finite variance σ_j^2 and d.f. F_j. Suppose that $B_N^2 = \sum_{j=1}^N \sigma_j^2$ satisfies the following condition, called* asymptotic negligibility condition,

$$\lim_{N \to \infty} \max_{1 \le j \le N} \frac{\sigma_j^2}{B_N^2} = 0.$$

If Φ denotes the d.f. of a standard normal, then

$$\lim_{N \to \infty} \Pr \left\{ \frac{\sum_{j=1}^N (L_j - \mu_j)}{B_N} \le z \right\} = \Phi(z)$$

if and only if the following condition, called Lindeberg condition,

$$\lim_{N\to\infty} \frac{1}{B_N^2} \sum_{j=1}^{N} \int_{\{|z-\mu_j|>\epsilon B_N\}} (z-\mu_j)^2 \, dF_j(z) = 0$$

is satisfied for every $\epsilon > 0$.

Proof. See Feller (1971), Section VIII.4. □

The Lindeberg-Feller CLT implies that, for large N, the distribution of L is well approximated by a normal distribution with mean $\sum_{j=1}^{N} \mu_j$ and variance B_N^2. The asymptotic negligibility condition requires that none of the individual variances σ_j^2 dominates the variance B_N^2 of L. Notice that the Lindeberg condition completely characterizes the conditions of asymptotic normality and asymptotic negligibility.

The Lindeberg-Feller CLT relies on independence, but it may be generalized to allow for restricted patterns of correlation. In general, the possibility of establishing a CLT implies that the distribution of the total loss L should look more "normal" than the distribution of potential losses for an individual loan. How large N must be for the distribution of the portfolio loss to look approximately normal, is a delicate issue that does not have a clear-cut answer.

4 Practitioners' Tools

4.1 Value-at-Risk (VaR) approach

The *Value-at-Risk (VaR)* is conventionally defined as the value of potential losses that will not be exceeded in more than a given fraction of possible events. This fraction, expressed in percentage terms, is called the "tolerance level".

Expected loss Is a loss that does not exceed the average loss.

Unexpected loss Is a loss that ranges between the average loss and the VaR. By definition, the probability of expected or unexpected loss is equal to one minus the tolerance level.

Exceptional loss Is a loss in excess of the VaR. By definition, the probability of exceptional loss is equal to the tolerance level.

4.1.1 VaR and quantiles

The statement that VaR = 100 at the tolerance level of 5 percent means that the chances of future losses exceeding 100 are equal to 5 percent. If losses are denoted by L, then the above statement is equivalent to

$$.05 = \Pr\{L > 100\} = 1 - \Pr\{L \leq 100\},$$

that is,

$$.95 = \Pr\{L \leq 100\},$$

which shows that the VaR is the .95th quantile of the distribution of the losses. More generally, the VaR at the tolerance level of $100 \times (1-u)$ percent is equal to the uth quantile $Q_L(u)$ of the distribution of the losses.

Because the loss is the difference $L = Y_0 - Y$ between the initial value and the end-of-period value, its d.f. is equal to

$$F_L(l) = \Pr\{L \leq l\} = \Pr\{Y \geq Y_0 - l\} = 1 - F(Y_0 - l).$$

This implies that

$$u = F_L(Q_L(u)) = 1 - F(Y_0 - Q_L(u)),$$

from which

$$1 - u = F(Q(u)) = F(Y_0 - Q_L(u)).$$

Hence

$$Q_L(u) = Y_0 - Q(1-u),$$

that is, the VaR at the tolerance level of $100 \times (1-u)$ percent may equivalently be expressed as the difference between Y_0 and the $(1-u)$th quantile of the distribution of the end-of-period value Y.

An important drawback of VaR is the fact that it is not additive. To see this, consider a portfolio consisting of two loans and let the r.v.s L_1 and L_2 represent their loss volatility. VaR is not additive because the quantile function of $L = L_1 + L_2$ is different from the sum of the quantile functions of L_1 and L_2.

4.1.2 A motivation for VaR

To provide motivation and additional insight for VaR, let the r.v. $L = Y - Y_0$ represent the uncertain end-of-period loss on a loan and consider the problem of how much capital should be set aside to absorbe potential losses.

Let the negative utility when c is the amount of capital set aside and l is a particular realization of L be represented by the number $q(l-c)$, where q is a nonnegative function which satisfies the following properties:

i. $q(0) = 0$;

ii. if $0 < z < z'$, then

$$q(0) \leq q(z) \leq q(z'), \qquad q(0) \leq q(-z) \leq q(-z');$$

iii. $q\colon \Re \to \Re_+$ is integrable with respect to the distribution of L.

The first condition is only an innocuous normalization. The second requires q to be nondecreasing for $z > 0$ and nonincreasing for $z < 0$. The third condition is satisfied if q is bounded from above, but otherwise restricts the class of problems that may be considered. Notice that we do not require q to be continuous, nor convex, nor symmetric, nor differentiable.

For a given c, the (negative) utility $q(L-c)$ is a transformation of the random variable L. Since $L-c$ is random, we consider its expectation

$$r(c) = E[q(L-c)]$$

and look for a c that makes $r(c)$ as small as possible. Clearly, the optimal choice of c depends on both the distribution of losses and the particular utility function q adopted.

Suppose that L has mean μ, median ζ and finite variance σ^2. If $q(z) = z^2$, then the optimal choice is $c^* = \mu$, with $r(c^*) = \sigma^2$. On the other hand, if $q(z) = |z|$, then the optimal choice is $c^* = \zeta$, with $r(c^*) = E(|L-\zeta|)$.

The main problem with the use of quadratic and absolute utility functions in our context is that they are symmetric about zero, which corresponds to the unrealistic assumption that the bank only cares about the size and not the sign of the difference $L - c$. In fact, the case when $L > c$ has very different implications for the bank than the case when $L < c$.

To incorporate the very different consequences of positive and negative deviations of L from c, consider the asymmetric absolute utility function

$$\begin{aligned} q_u(z) &= [u\,1\{z \geq 0\} + (1-u)\,1\{z < 0\}]\,|z| \\ &= [u - 1\{z < 0\}]\,z, \qquad 0 < u < 1, \end{aligned}$$

where $1\{A\}$ is the indicator function of the event A. Unless $u = 1/2$, which corresponds to symmetric absolute utility, losses are now penalized differently depending on whether $l > c$ or $l < c$. When $u > 1/2$, losses in excess of c are penalized more heavily, and increasingly so as u increases. It can be shown that, in this case, the optimal choice of c is $c^* = Q_L(u)$, the uth quantile of L.

Notice that we obtain a different solution to the problem if we use instead the following asymmetric modification of quadratic utility

$$q_u(z) = [u\,1\{z \geq 0\} + (1-u)\,1\{z < 0\}]\,z^2.$$

If $f(l)$ denotes the density of L, then the optimal choice of c in this case can be shown to satisfy the relationship

$$\frac{\int_{c^*}^{\infty}(l - c^*)\,f(l)\,dl}{\int_{-\infty}^{c^*}(c^* - l)\,f(l)\,dl} = \frac{1-u}{u},$$

which defines the uth *expectile* of L (see Newey and Powell, 1987). Thus, expectiles may provide an alternative to quantiles for VaR calculations.

4.1.3 Sensitivity to the measure of volatility

Common practice of computing the VaR is

$$\text{VaR} = \text{loss volatility} \times \text{multiple of volatility},$$

where the multiple of volatility is a number that depends both on the assumptions about the loss distribution and the chosen tolerance level.

This practice is justified whenever the loss can be represented as

$$L = \mu + \sigma U, \qquad \sigma > 0,$$

where μ is a measure of center of the loss distribution (generally, $\mu = 0$), σ is measure of volatility and U is r.v. with a known distribution. In this case, the VaR at the tolerance level of u percent is equal to

$$Q_L(u) = \mu + Q_U(u)\,\sigma,$$

where $Q_U(u)$ is the uth quantile of the r.v. U. Of course, using 1.96 or 2.33 as multiples of volatility may not be appropriate if the distribution of U is not normal, as it is typically the case whith credit-related losses.

The measure of volatility need not be the standard error of the loss. Alternative measures, especially useful in the case of financial time series, include a weighted average of past observations with exponentially declining weights

$$\sigma_t^2 = (1-\lambda)\sum_{i=0}^{n} \lambda^i (Y_{t-i} - \mu)^2 \approx \lambda \sigma_{t-1}^2 + (1-\lambda)(Y_{t-1} - \mu)2,$$

and ARCH, GARCH or EGARCH models of volatility. On the sensitivity of VaR to different measures of volatility, see Drudi, Generale and Majnoni (1997).

4.1.4 Choice of tolerance level

The choice of the tolerance level is an open issue. The Basle Committee on Banking Supervision recommends 1 percent, but it is intuitively clear that the tolerance level for day-to-day operations may be higher than the case when solvency issues are involved.

An important question is how sensitive is the computed VaR to the choice of tolerance level. Recall that VaR at the tolerance level of $100(1-u)$ percent is equal to the uth quantile $Q_L(u)$ of the loss distribution.

From the relationship $F_L(Q_L(u)) = u$ we get

$$Q_L'(u) = \frac{1}{f_L(Q_L(u))},$$

where $f_L = F_L'$ denotes the density of the loss distribution. Viewed as a function of u, this is known in the statistical literature as *sparsity* or *quantile-density* function [see Parzen (1979)]. The quantile-density function is well defined whenever the loss density is strictly positive and is strictly positive whenever the density is bounded.

4.2 CreditMetrics^{TM} and CreditRisk$^+$

We now introduce two recent models for evaluating credit risk which have gained considerable attention among practitioners. CreditMetrics^{TM} is a portfolio model for evaluating the risk arising from credit quality changes caused by upgrades, downgrades and defaults. The methodogy has been developed by J.P. Morgan and a set of co-sponsors consisting of Bank of America, Bank of Montreal, BZW, Deutsche Morgan, KMV Corporation, Swiss Bank Corporation, and Union Bank of Switzerland.

CreditRisk$^+$ is instead a model of credit risk that takes into account the volatility of default rates. It has been developed by Credit Suisse Financial Products.

By providing estimates of the loss distribution from a portfolio of exposures, both models may be used for VaR and CaR calculations. Neither model, however, makes an attempt at pricing the risk of a loan or a portfolio of loans.

4.2.1 CreditMetrics^{TM}

The starting point of CreditMetrics^{TM} (CM) is the fact that "migrations", that is, changes in credit quality (due to credit upgrades or downgrades, not only defaults) cause changes in the value of a loan which are ignored by book value accounting.

CM tries to assess individual and portfolio VaR due to credit changes by proceeding in four main steps:

i. computation of the exposure profile of each loan;

ii. computation of the volatility of value due to credit quality changes;

iii. estimating credit quality correlations across loans;

iv. calculating the distribution of values for the portfolio.

The first step is necessary because, while some loans have a fixed exposure amount, others create exposures which can vary.

The second step involves:

- estimating credit quality "migrations" (transition matrix between credit ratings);
- estimating changes in value upon credit quality migration;
- computing the distribution of values for a single instrument or some summary of this distribution (mean, standard deviation).

Estimating credit quality correlations (step 3) need not be an easy task. In general, independence is too strong an assumption. Correlation is likely to be present because rating outcomes on different loans are not independent of each other, being affected in part by the same economic factors.

CM argues that "empirically, correlation data are the most complex and potentially controversial element in credit portfolio modeling". Typically, the alternative is between using data that are sparse or of poor quality and employing strong assumptions.

The approach of Merton (1974) creates a link between the underlying value of a firm and its credit rating, and ultimately corresponds to an ordered logit or probit model. This allows the joint probability for two firms to be built from a knowledge of the correlation between two firms's asset values, which are proxied by their equity prices.

To simplify things, CM estimates a correlation matrix that is structured into blocks corresponding to relatively homogeneous groups of loans in the portfolio (e.g. grouping may correspond to industry/country combinations).

Finally, the last step consists of calculating the distribution of values for the portfolio. Unless the number of loans N and credit states S is small, it is unfeasible to consider all possible portfolio states in order to obtain the distribution of values for the portfolio. Instead, a sample of states is selected at random, and the distribution of values for the portfolio is estimated from this sample. If the sample is chosen at random and is large enough, then this simulation approach will produce distributions that are arbitrarily close to the actual one.

4.2.2 CreditRisk$^+$

Credit Suisse Financial Products' CreditRisk$^+$ (CR+) tries to compute the distribution of default losses by proceeding in two steps:

i. computation of the distribution of default events in a given time period by modeling default rates as continuous r.v.s;

ii. computation of the distribution of different exposures.

Treating default rates as continuous r.v.s allows CR+ to incorporate uncertainty in the level of the default rate. Allowing for default rate volatility does not change the expected loss but give rise to loss distributions with fat tails.

The second step is necessary because the distribution of losses differs from the one of defaults as the loss in a given default depends on the exposure. In the CR+ methodology, exposures are net of expected recovery rates and are discretized into exposure bands.

The main advantage over CM is the low data requirement and the computational efficiency due to the availability of closed form solutions. The main drawbacks are its more limited scope and the restrictive nature of some of its assumptions.

4.3 Two basic models

We now present two basic models that try to capture correlation in credit grade changes of different loans. A good grasp of these two alternative models is essential in order to understand the main differences between CM and CR+.

4.3.1 The ordered threshold model

In this model, which represents the basis of CM, credit events are driven by movements in the underlying value of an obligor, represented by a latent r.v. Y^*. Its key feature is that credit grade changes are correlated across obligors because of correlation in their latent values due to the common influence of macroeconomic, sectoral or country specific factors.

We generalize the model in Section 3.2.1 by assuming that, besides default, there are $S-1$ other credit states, corresponding to increasing credit grade. The indicator D now takes $S \geq 2$ values, one for each possible end-of-period credit state, with $D = 0$ in case of default. To keep notation as simple as possible, we do not make explicit the dependence of credit state probabilities on a vector X of observable characteristics, which may include the current credit state.

Credit states are defined in terms of S cut-off values

$$C_1 < C_2 < \ldots < C_S,$$

with $C_S = \infty$. The loan is in state 0 (default) whenever $Y^* < C_1$, and is in state $d = 1, \ldots, S-1$ whenever Y^* falls in the interval $[C_d, C_{d+1})$. Thus,

$$\pi_d = \Pr\{D = d\} = \begin{cases} \Pr\{Y^* < C_0\} = F(C_0), & \text{if } d = 0, \\ \Pr\{C_d \leq Y^* < C_{d+1}\} = F(C_{d+1}) - F(C_d), & \text{if } d = 1, \ldots, S-1, \end{cases}$$

with $F(C_S) = 1$.

Important practical problems are:

- the choice of F.
- the choice of the number and position of the cut-offs $C_0, \ldots, C_{S-1}$.

Example 3 If the latent r.v. Y^* is distributed as normal with mean μ and variance σ^2, then

$$\pi_d = \Pr\{C_d \leq \mu + \sigma U < C_{d+1}\} = \Phi\left(\frac{C_{d+1} - \mu}{\sigma}\right) - \Phi\left(\frac{C_d - \mu}{\sigma}\right),$$

for $d = 1, \ldots, S-1$, which is just the *ordered probit* model. We obtain the *ordered logit* model if the distribution of U is instead assumed to be logistic.

The use of symmetric unimodal distributions for Y^* is consistent with the observed evidence that transitions are low-probability events and that transition probabilities are inversely relate to the distance from the current state.

□

Now let us go back to the special case of only two end-of-period states (default and no default), but assume that there are $N \geq 1$ obligors indexed by $j = 1, \ldots, N$. Let D_j and Y_j^* be, respectively, a binary indicator of default and the latent end-of-period value of the jth obligor. Default occurs whenever Y_j^* falls below some cut-off value C_j.

The vector $D = (D_1, \ldots, D_N)$ of default indicators takes 2^N possible values, whith probabilities that depend on the N-vector $C = (C_1, \ldots, C_N)$ of cut-off values and the joint distribution of $Y^* = (Y_1, \ldots, Y_N)$, the random N-vector of latent values. Thus, for example, the probability that all loans default is

$$\Pr\{D_1 = 0, \ldots, D_N = 0\} = \Pr\{Y_1^* < C_1, \ldots, Y_N^* < C_N\},$$

whereas the probability that no loan defaults is equal to

$$\Pr\{D_1 = 1, \ldots, D_N = 1\} = \Pr\{Y_1^* \geq C_1, \ldots, Y_N^* \geq C_N\}. \tag{4.48}$$

An important special case is when the obligors' values are mutually independent, for then

$$\Pr\{D_1 = d_1, \ldots, D_N = d_N\} = \prod_{j=1}^{N} \Pr\{D_j = d_j\},$$

where d_j is either 0 or 1, and $\Pr\{D_j = d_j\} = F_j(C_j)^{1-d_j}[1 - F_j(C_j)]^{d_j}$.

In the general case, however, calculating (4.48) requires evaluating an N-dimensional integral. In particular, if the joint distribution of Y^* is normal, then numerical integration is only feasible when $N \leq 4$. Unless the joint distribution of Y^* is restricted in some way, these N-variate integrals do not have a closed form solution and must be approximated by Monte Carlo methods.

Example 4 Suppose that there are only two obligors ($N = 2$) and let $Y^* = (Y_1^*, Y_2^*)$ have a bivariate normal distribution with mean $\mu = (\mu_0, \mu_1)$ and variance

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}.$$

Then $D = (D_1, D_2)$ can take four values with the following probabilities

$$
\begin{aligned}
\Pr\{D = (0,0)\} &= \int_{-\infty}^{C_1} \int_{-\infty}^{C_2} f(y_1, y_2)\, dy_2\, dy_1, \\
\Pr\{D = (1,0)\} &= \int_{C_1}^{\infty} \int_{-\infty}^{C_2} f(y_1, y_2)\, dy_2\, dy_1, \\
\Pr\{D = (1,0)\} &= \int_{-\infty}^{C_1} \int_{C_2}^{\infty} f(y_1, y_2)\, dy_2\, dy_1, \\
\Pr\{D = (1,1)\} &= \int_{C_1}^{\infty} \int_{C_2}^{\infty} f(y_1, y_2)\, dy_2\, dy_1,
\end{aligned}
$$

where $f(y_1, y_2)$ is the bivariate $\mathcal{N}(\mu, \Sigma)$ density.

Now suppose that Y_1^* and Y_2^* are conditionally uncorrelated given a vector X of risk factors. Then we may write

$$
\begin{aligned}
Y_1^* &= X^\top \gamma_1 + \sigma_1 U_1, \\
Y_2^* &= X^\top \gamma_2 + \sigma_2 U_2,
\end{aligned}
$$

where U_1 and U_2 are uncorrelated standard normal, and the previous bivariate integrals simplify to

$$
\begin{aligned}
\Pr\{D = (0,0)\} &= \Phi(C_1^*(X))\, \Phi(C_2^*(X)), \\
\Pr\{D = (1,0)\} &= [1 - \Phi(C_1^*(X))]\, \Phi(C_2^*(X)), \\
\Pr\{D = (1,0)\} &= \Phi(C_1^*(X))\, [1 - \Phi(C_2^*(X))], \\
\Pr\{D = (1,1)\} &= [1 - \Phi(C_1^*(X))]\, [1 - \Phi(C_2^*(X))],
\end{aligned}
$$

where

$$
C_j^*(X) = \frac{C_j - X^\top \gamma_j}{\sigma_j}, \qquad j = 1, 2.
$$

and $\Phi(u)$ is the d.f. of a standard normal. □

The case with $N \geq 1$ obligors and $S \geq 2$ end-of-period states is a simple generalization of Example 4. The vector $D = (D_1, \dots, D_N)$ of default indicators now takes S^N possible values, whith probabilities that depend on an $N \times S$ matrix C of cut-off values and the joint distribution of $Y^* = (Y_1, \dots, Y_N)$. The jth row $C_j = (C_{j1}, \dots, C_{jS})$ of C displays the cut-off values for the jth obligor.

If $Y_1^*, \dots, Y_N^*$ are conditionally uncorrelated given a K-vector $X = (X_1, \dots, X_K)$ of risk factors, then a convenient representation of Y_j^* is

$$
Y_j^* = \mu_j + \sigma_j U_j, \qquad j = 1, \dots, N,
$$

where μ_j is a systematic component, correlated across loans, which depends on the vector of risk-factors, $\sigma_j > 0$ is a scale parameter, and U_j is a purely idiosyncratic component, uncorrelated across loans. A typical specification of μ_j is

$$\mu_j = \sum_{k=1}^{K} X_k \gamma_{jk} = X^\top \gamma_j,$$

where $\gamma_j = (\gamma_{j1}, \ldots, \gamma_{jK})$ is a vector of "factor loadings". Turning to the cut-offs, a convenient assumption is that they depend on the loan only through its initial credit grade G_j, that is, $C_{jd} = C_d(G_j)$. Under these two specifications, the jth loan is in state 0 (default) at the end of the period whenever $Y_j^* < C_{j1}$, that is, $D_j = 0$ whenever

$$U_j < \frac{C_1(G_j) - X^\top \gamma_j}{\sigma_j},$$

and is in state $d = 1, \ldots, S-1$ whenever Y_j^* falls in the interval $[C_{jd}, C_{j,d+1})$, that is, $D_j = d$ whenever

$$\frac{C_d(G_j) - X^\top \gamma_j}{\sigma_j} \leq U_j < \frac{C_{d+1}(G_j) - X^\top \gamma_j}{\sigma_j}.$$

Notice that, although we managed to avoid the need of computating an N-dimensional integrals, the set of probabilities implied by the model is large even for moderate values of S and N. For example, with 4 end-of-period credit states ($S = 4$) and 10 obligors ($N = 10$), the number of possible values of the vector D is over 1 million.

4.3.2 The Poisson model

The model outlined in Section 4.3.1 is computationally demanding. The model present here, which is at the heart of CR+, tries to reduce the computational burden by relying on closed-form solutions. As we shall see, however, nothing comes for free. A first restrictive assumption is that there are only two end-of-period states, default and no-default.

For convenience, we depart from our previous notation by letting the default indicator D take value 0 if no default occurs and value 1 otherwise.

The case of a single loan is essentially the same as in the threshold model. The main difference is that the dependence of the default probability π on the initial credit grade $G = g$ and the vector of risk factors $X = x$ is postulated as

$$\pi(g, x) = \bar{\pi}(g) \sum_{k=1}^{K} x_k \gamma_k = \bar{\pi}(g) \cdot x^\top \gamma. \tag{4.49}$$

No attempt is made at modelling the causes of default. If the components of X are nonnegative r.v.s with unit mean, then

$$E[\pi(g,X)] = \bar{\pi}(g) \sum_k \gamma_k.$$

Hence $E[\pi(g,X)] = \bar{\pi}(g)$ provided that the elements of the vector β add up to one, in which case $\bar{\pi}(g)$ may be interpreted as the average default probability of an obligor with credit grade $G = g$.

Notice that specification (4.49) implies that the ratio

$$\frac{\pi(g',x)}{\pi(g,x)} = \frac{\bar{\pi}(g')}{\bar{\pi}(g)}$$

does not depend on x, whereas the ratio

$$\frac{\pi(g,x')}{\pi(g,x)} = (x'-x)^\top \gamma$$

does not depend on the credit grade g.

Because the r.v. D has a Bernoulli distribution, its conditional probability generating function (p.g.f.) given $G = g$ and $X = x$ is

$$\begin{aligned} M(z \mid g,x) = E(z^D \mid G=g, X=x) &= \sum_{d=0}^{\infty} z^d \Pr\{D = d \mid G = g, X = x\} \\ &= 1 - \pi(g,x) + \pi(g,x)z = 1 + \pi(g,x)(z-1). \end{aligned}$$

Because $\ln(1+y) = y$ for small y, we get

$$M(z \mid g,x) = \exp \ln[1 + \pi(g,x)(z-1)] \approx \exp[\pi(g,x)(z-1)], \qquad (4.50)$$

which corresponds to the p.g.f. of a Poisson distribution with parameter (arrival rate) $\pi(g,x)$.

Now consider the case of $N \geq 1$ obligors, each with default indicator D_j. From (4.50), the conditional p.g.f. of D_j given G_j and $X_j = x$ is

$$M_j(x) \approx \exp[\pi_j(x)(z-1)], \qquad j = 1, \ldots, N,$$

where $\pi_j(x) = \bar{\pi}_j \cdot x^\top \gamma_j$ and $\bar{\pi}_j = \bar{\pi}(G_j)$.

Our key assumption is that, conditionally on the vector of risk factors X, default events are independently distributed across obligors. This assumption is similar to the one made in Section 4.3.1 and enables us to exploit the following properties of p.g.f.s:

i. The p.g.f. of the sum of independent random variables is equal to the product of their p.d.f.s.

ii. If $M(z \mid x)$ is the conditional p.g.f. given $X = x$, the unconditional p.g.f. is equal to $E[M(z \mid X)]$.

The first property implies that, conditional on $X = x$, the p.g.f. of the total number of defaults $\sum_{j=1}^{N} D_j$ is

$$M(z \mid x) = \prod_{j=1}^{N} M_j(z \mid x) \approx \exp[\pi(x)(z-1)],$$

where $\pi(x) = \sum_{j=1}^{N} \pi_j(x)$. Notice that $M(z \mid x)$ is approximately equal to the p.g.f. of a Poisson distribution with parameter $\pi(x)$.

The second properties implies that

$$M(z) = \int \prod_{j=1}^{N} M_j(z \mid x)\, h(x)\, dx, \tag{4.51}$$

where $h(x)$ denotes the joint density of X. To get a closed form expression for $M(z)$, it is convenient to assume that the components of the vector X are independent gamma r.v.s.

Recall that the gamma is a two-parameter family of distributions with density of the form

$$f(u) = \frac{1}{\beta^\alpha \Gamma(\alpha)}\, e^{-u/\beta} u^{\alpha-1}, \qquad u \geq 0,$$

where $\Gamma(\alpha) = \int_0^\infty e^{-u} u^{\alpha-1}\, du$ is the gamma function, and the parameters α and β are related to the mean μ and the variance σ^2 of the distribution through

$$\alpha\beta = \mu, \qquad \alpha\beta^2 = \sigma^2.$$

Under the assumption that the kth component of X has a gamma distribution with unit mean and variance equal to σ_k^2, evaluating (4.51) yields

$$\begin{aligned} M(z) &\approx \prod_{k=1}^{K} \int_0^\infty \exp[c_k x_k (z-1)]\, h_k(x_k)\, dx_k \\ &= \prod_{k=1}^{K} \left(\frac{1-\delta_k}{1-\delta_k z} \right)^{1/\sigma_k^2}, \end{aligned}$$

where $c_k = \sum_{j=1}^{N} \bar{\pi}_j \gamma_{jk}$ and

$$\delta_k = \frac{\sigma_k^2 c_k}{1 + \sigma_k^2 c_k}.$$

Notice that

$$\left(\frac{1-\delta_k}{1-\delta_k z}\right)^{1/\sigma_k^2}$$

is the p.g.f. of a negative binomial distribution. Thus, the total number of defaults is approximately distributed as the sum of K independent negative binomial variates.

It is worth summarizing the key assumptions behind this result:

i. The "proportional hazard" specification of $\pi_j(x)$.

ii. The assumption that, conditionally on X, default events are independently distributed across obligors.

iii. The use of the approximation (4.50).

iv. The assumption that the components of the random vector X are independently distributed as gamma variates with unit mean.

Of these assumptions, 2. and 4. are the most restrictive.

5 Extensions and Applications of Pricing

5.1 Term Structure of Credit spreads

Default risk is not constant. It varies over time and, in any given period, there is a time structure to it.

Consider our old pricing equation

$$v(t,T,n) = \frac{1}{R^{T-t}}\left(\delta[q + q(1-q) + \dots (1-q)^{T-t-1}] + (1-q)^{T-t}\right) \tag{5.52}$$

The risk spread is the difference between the price when $q = 0$ and the case just considered

$$\frac{v(t,T,n)}{R^{-(T-t)}} = \delta[q + q(1-q) + \dots (1-q)^{T-t-1}] + (1-q)^{T-t} \tag{5.53}$$

Notice the following, important, points

> The margin depends upon maturity, the closer the maturity the smaller the margin.

In the extreme case $\delta = 0$, the ratio is $(1-q)^{T-t}$. Hence the margin gives an implicit estimate of the istantaneous default probability

$$q = 1 - \left(\frac{1 + Y(t,T)}{1 + Y_d(t,T)}\right)^{\frac{1}{T-t}}$$

where $1 + Y_d(t,T)$ and $1 + Y(t,T) = R$ are the rates, compounded over the period from t to T, for a risky and risk free security respectively. $1 + Y(t,T) = R$. Hence

$$R_d = \frac{R}{1-q}.$$

It is easy to generalize this method for computing default's probabilities to the case in which the probability of defaulting at time t, conditional on not having defaulted yet, is different from the probability of defaulting at time $t+1$, again conditional upon getting safely there. To do this, obviously, one needs more than just one risk free rate (or price of risk free asset) and one risky rate.

In fact, as argued in class, one needs a pair of prices for each future date t at which the conditional default probability must be computed. For each maturity we need the price of a risk free zero coupon bond and a risky zero coupon bond, both maturing at that date. Then the previous formula can be applied recursively to extract the whole term structure of default probabilities $\{q_t, q_{t+1}, \dots, q_T\}$.

5.2 Continuous Time Pricing

Simplest case

A security is expected to pay X next period. With probability h it defaults, and with probability $1-h$ it does not. At defaults it pays only $1-L$ of face value X. Hence, expected payment is

$$x = X[h(1-L)+h]$$

What's the implicit rate of interest $R_d = 1 + r_d$ for such a security, as a function of the risk free rate $(1+r) = R$? NA requires

$$\frac{1}{1+R} = \frac{1}{R_d}[h(1-L)+(1-h)] = \frac{1}{R_d}[1-hL] \tag{5.54}$$

$$r_d = \frac{r+hL}{1-hL}. \tag{5.55}$$

The extension to a multi period setting, with future dates $t_1, t_2, \ldots, t_n$, at which borrower must pay C_k, per $k = 1, \ldots, n$, given probabilities p_k can be derived with our methods.

Start from value of loan when there is no default risk:

$$P = \sum_k C_k e^{-rt_k} \tag{5.56}$$

whereas, with risk of default, we get:

$$P_d = \sum_k p_k C_k e^{-rt_k}. \tag{5.57}$$

Then we can define r_d implicitely, as

$$P_d = \sum_k C_k e^{-r_d t_k} \tag{5.58}$$

Notice, obviously, that in this intuitive explanation we are making the assumption of knowing all the true probabilities of default at all future dates. It should be easy at this point, though, to replace the true probabilities with the martingale equivalent probabilities implied by no arbitrage and proceed likewise. In general, in fact, we want to think of r_d as equal to the risk free rate r plus a premium which is due to the default probability implicit in the risk neutral pseudo probabilities.

Poisson Rate

If the arrival rate of the default event follows a Poisson process with parameter λ, the probability of being alive at t is

$$e^{-\lambda t} \tag{5.59}$$

In this case the value of r_d is easy to compute:

$$r_d = r + \lambda \tag{5.60}$$

This is nice, simple and also intuitive. In particular, given that

$$1 - e^{-\lambda} \simeq \lambda, \tag{5.61}$$

the spread between r_d and r is a direct and simple measure of the probability of defaulting, per unit of time.

Back to continuous time

Let us take the limit and consider r_d, r and h as istantaneous rates ver infinitesimal time units, $r_d\Delta t$, $r\Delta t$, $h\Delta t$. Then

$$r_d\Delta t = \frac{r\Delta t + h\Delta t L}{1 - h\Delta t L}. \tag{5.62}$$

Dividing by Δt,

$$r_d = \frac{r + htL}{1 - h\Delta t L}. \tag{5.63}$$

we obtain a formula that converges, in the limit, to the continuous time interest rate for defaultable contracts

$$r_d = r + hL. \tag{5.64}$$

5.3 Continuous Time Markov Processes

Let the state of the system be Y_t. This is a stochastic process. At time T a state dependent value $g(Y_T)$ is to be paid.

The instantaneous interest rate also follows a state dependent process

$$\rho_t = \rho(Y_t).$$

If the interest rate were deterministic, the value of the security would be

$$V = e^{-\rho(T-t)} g(Y_T)$$

which generalizes to the current situation as

$$V(y_t, t) = E[e^{(\int_t^T -\rho(Y_s)ds)} g(Y_T) \mid Y_t]. \tag{5.65}$$

Assume the state process follows

$$dY_t = \mu(Y_t)dt + \sigma(Y_t)dB_t,$$

then the value of the security must solve the stochastic differential equation

$$\begin{aligned} 0 = & -\rho(y)V(y,t) + V_t(y,t) + V_y(y,t)\mu(y) \\ & +1/2 \text{ trace}[\sigma(y,t)V_{yy}(y,t)\sigma(y,t)] \end{aligned} \tag{5.66}$$

and the boundary value restriction

$$V(y,T) = g(y) \text{ for all } y. \tag{5.67}$$

5.4 Modified interest rate

The formula

$$r_d = r + hL$$

can be extended as follows.

Let $r(t)$ be the time t risk free rate. Assume the default process is such that the default time T can be represented as follows

$$\Lambda(t) = \mathbf{1}_{\{t \geq T\}}.$$

In other words, $\Lambda(t)$ is a process equal to zero before default and to 1 after the default event occurrs. This implies that the instantaneous default rate $h(t)$ satisfies:

$$d\Lambda(t) = h(t)dt + dM(t)$$

where M is a martingale. Assume the recovery or repayment rate, in the event of default, is $L(t) \in [0,1]$, then one can show that the modified interest rate is

$$r_d(t) = r(t) + h(t)L(t).$$

With this modified interest rate one can compute the current value of all payment processes (i.e. securities) that are exposed to a risk of default behaving according to the specified pair (h, L).

$$E\left[\exp\left(\int_0^t -r_d(s)ds\right)X\right] \tag{5.68}$$

Next we use the general results just derived, together with option pricing theory, to price different kinds of securities.

5.5 Options on spreads

We begin with a simple option on a spread, where the spread is the premium to be paid for the underlying security being exposed to default risk. In other words, this option is written on a security which has the following properties

i. pay a coupon every six months, and expires at T_m,

ii. its price is based on a spread over a risk free security with the same maturity T_m and equal coupon flow.

Let Z_t be the risk free return, and S_t the spread.

The price of the underlying security is then written as

$$p(t, S_t + Z_t).$$

We assume we observe such a price. The option we use has the following characteristics. It gives the right of selling the security at a price such that you earn a spread S^L over Z_T at time T. Recall that the higher the implied spread (i.e. the higher the probability of default has become for the underlying security) the lower will be the market value of the security. If, at maturity, the spread is such

$$S^L + Z_T \geq S_T + Z_T$$

the option expires valueless; otherwise it is worth

$$p(T, S^L + Z_T) - p(T, S_T + Z_T).$$

Often, when options on spread are sold, seller may add a threshold to cover themselves from extremely high volatility. In this case, when the spread reaches a level $S^H > S^L$ at any time $t < T$, the option pays

$$p(t, S^L + Z_t) - p(t, S^t + Z_t) = p(t, S^L + Z_t) - p(t, S^H + Z_t),$$

and automatically expires in that moment.

In these circumstances the option has exercise values that change according to three different situations. Hence, the value X for such option will be:

$$X = \max[p(\tau, S^L + Z_\tau) - p(\tau, \min(S_\tau, S^H) + Z_\tau), 0] \tag{5.69}$$

where

$$\tau = \min(T, \min\{t : S_t \geq S^H\}).$$

This is the terminal condition we need. Now all you need to do is to repeat the procedure described above. Given a stochastic process for the istantaneous default probability, use Ito's calculus and this boundary condition. This, together with the formula given earlier in 5.68, where the payoff X is given by 5.69, provides the pricing equation.

5.6 Further Applications

5.6.1 Swaps

An (*interest swap*) between parties A and B specifies

i. A pays B for a certain period of time, a quantity equal to a certain fixed interest rate over X units of value;

ii. B pays A for a certain period of time, a quantity equal to a certain variable interest rate over X units of value; such variable rate is equal to a market rate plus a prespecified spread;

Let's compute the value of such swap for one of the two parties. We take as a reference the party that

i. pays floating and receives fix;

ii. is not at risk of defaulting;

iii. the other part is at risk of defaulting;

iv. the contract lasts for two periods.

Let r be the fixed and $r(t)$ the variable rates, for $t, t = 1, 2$; denote with $p(t, T)$ the value at t of a unit payment at T. The present value of the payment expected in the first period (period $t = 0$) by our party is

$$\begin{aligned} F(0,1) &\equiv \frac{r(1)}{1+r(1)} \\ &= 1 - p(0,1) \end{aligned} \tag{5.70}$$

and that of the second is

$$\begin{aligned} F(0,2) &\equiv \frac{r(2)}{(1+r(1))(1+r(2))} \\ &= p(0,1) - p(0,2) \end{aligned} \tag{5.71}$$

Hence, the value to our party of the exchange in the first period is

$$\begin{aligned} v_s(0,1) &\equiv \frac{r - r(1)}{1+r(1)} \\ &= rp(0,1) - F(0,1) \end{aligned} \tag{5.72}$$

and in the second

$$v_s(0,2) \equiv \frac{r - r(2)}{(1+r(1))(1+r(2))}$$
$$= rp(0,2) - F(0,2). \quad (5.73)$$

From which we conclude that

$$[rp(0,1) - F(0,1)](q\delta + (1-q)) + [rp(0,2) - F(0,2)](q\delta + (1-q)q\delta + (1-q)^2) \quad (5.74)$$

is the value of the swap for our side.

Often $\delta = 0$, and so we simplify to

$$[rp(0,1) - F(0,1)](1-q) + [rp(0,2) - F(0,2)](1-q)^2 \quad (5.75)$$

5.6.2 European Options

An *European call option*, expiring at T, with strike price K, on a security which is exposed to default risk. Go back to the old notation in which the state had the structure $(s,x) \in \{h,l\} \times \{n,f\}$. Do backward induction beginning again in the last, second, period:

$$C(1,(s,x)) = \max[v(1,2,(s,x)) - K, 0] \quad (5.76)$$

hence in the first period

$$C(0) = (1-q)[\pi C(1,(h,n)) + (1-\pi)C(1,(l,n))] \quad (5.77)$$
$$+q[\pi C(1,(h,f)) + (1-\pi)C(1,(l,f))].$$

The notation here means that q is the default probability and π is the probability of the h aggregate state of the world. Both are risk neutral probabilities, obviously.

5.7 Coupon-paying securities

The security now pays a predetermined coupon y_t every time up to maturity, i.e. it pays the sequence

$$y = (y_1, \ldots, y_t, \ldots, y_T, 0, 0, \ldots). \quad (5.78)$$

As usual, the two-dimensional state vector $\omega_t = (s_t, x_t)$ describes the macro variable, interest rate say, and the borrower individual position respect to default.

Again, we have the following end-values

$$v(T,X,(s,n)) = y_T, v(T,X,(s,f)) = \delta y_T,$$

for all states s.

From here we proceed backward. At any intermediate period we find that

$$v(t, X, (s, n)) = y_t + \frac{1}{R(s_t)} \tag{5.79}$$

$$\sum_{(s_{t+1},\ldots,s_T)} \{ \frac{q(s_{t+1}) \ldots q(s_{T-1})}{R(s_{t+1}) \ldots R(s_{T-1})} \left[\sum_{\tau=t}^{T-1} q(1-q)^{\tau-t} [(y_t + \ldots + y_{\tau-1}) + \delta(y_\tau + \ldots + y_T)] \right]$$

$$+(1-q)^{T-t}(y_t + \ldots + y_T).\}$$

The formula 5.79 is long, but easy to interpret if we go through the following list of components for the full value of the security.

i. payment at each t, y_t;

ii. sum of future payments, discounted at the compound interests $R(s_{t+1}) \ldots R(s_{T-1})$ which occurr with probabilities $q(s_{t+1}) \ldots q(s_{T-1})$,

iii. the payments $y_t + \ldots + y_{\tau-1}$ are in full until default takes place at period τ, after which we receive a fraction δ of residual payments $y_\tau + \ldots + y_T$;

iv. and, finally, the probability of default taking place at τ is $q(1-q)^{\tau-t}$.

6 Describing Fancy Securities

Distinction between the credit risk of derivative instruments and securities that are credit derivatives. Here we quickly list fundamental and more popular types of credit derivatives.

6.0.1 Insurance Based Credit Derivatives

- Default Put e Digital Put
- Credit or Default Swap
- Total Return Swap
- First-to-Default Swap

Second generation credit derivatives include more sophisticated animals, such as, for example

- Put Options on Defaultable Bonds
- Credit Spread Put Options
- Index Swap
- Credit Swaps Dinamici

7 Valuation of credit derivatives

This is a fairly complicated topic, which we cannot address here in full. Here is a super-fast summary of the various steps and tools required to price an object with the following characteristics.

Assume either risk neutrality or that all probabilities used in the following are martingale equivalent ones.

Risk Free Term Structure

This is represented by the price at t of zero coupon bonds maturing at T. Call it $P(t,T)$.

Defaultable Bond Prices

Assume these are also zero coupon and indicate with $D(t,T)$ their value at t. Again, we assume this is an interpolated term-structure from market data on traded securities observed at t.

Default Process

From t until termination at T, default event can be realized at intervals τ_i, $i \in N$. This is an increasing sequence of stopping time. The latter are induced by a Cox process, characterized by a non negative intensity $h(t)$. Notice that the event default, i.e. the jump at some τ_i is independent from past default events. This contradicts evidence, as evidence shows that default behaviors have memory and past default history helps in predicting future. In this model, instead, the time of default τ_i is modelled as inaccessible.

Notice that we allow for many default events, i.e. once default hits you do not disappear. This allows the model to capture downgrading of security, partial suspensions, etc.

The intensity h can be stimated from historical defautl data for bonds of same risk class.

Recovery rates

At any stopping time τ_i associate a stochastic value q_i, which is the loss percentage.

Hence we have the following terminal value as fraction of market value at T

$$Q(T) = \Pi_{\tau_i \leq T}(1 - q_i) \tag{7.1}$$

Former results can be used to show that (time indeces omitted)

$$r' = r + hq \tag{7.2}$$

where is the default-adjusted rate and q is the expected value at the appropriate istant.

Let $R(t,T)$ be the discount factor for risky security, at t for horizon T, defined as

$$R(t,T) = E_t[\exp(-\int_t^T r'(s)ds)] \tag{7.3}$$

Use(7.1) and (7.3) to get the price at t of the default bond

$$D(t,T) = Q(t)R(t,T) \tag{7.4}$$

Recall we make the assumption that the default process and the risk free rate process are independent. This is strong and can be relaxed. This assumption, though, allows us to use (7.4) to compute survival probabilities at t for the risky bond. Indicate it with $S(t,T)$, it is the conditional probability at t of no default over $[t,T]$. It is

$$S(t,T) = \frac{1}{Q(t)}\frac{D(t,T)}{P(t,T)} \tag{7.5}$$

Given that $Q(0) = 1$, one has

$$S(0,T) = \frac{D(0,T)}{P(0,T)} \tag{7.6}$$

Work by Duffie, Lando, Madan, Schönbucher, Unal show that this model permits the explicit computation of first generation credit derivatives prices.

This model takes as inputs the following data, all measurable at $t = 0$:

(1) $P(0,T)$;

(2) Under $q = 1$, $D(0,T)$;

(3) $S(0,T)$.

Given this data, we have the following

Default Digital Option premium $X(0)$

$$X(0) = P(0,T) - D(0,T) \tag{7.7}$$

Default Swap rate c

$$c = \frac{\int_0^T D(0,t)h(0,t)dt}{\int_0^T D(0,t)dt} \tag{7.8}$$

7.0.2 Implicit Default Probabilities

This procedure requires four steps

i. Estimate V and σ_V.

ii. Given D, calculate the "default point".

iii. Compute expected value of V at horizon and compare it to value of debt payments due. This provides us with a "distance to default".

iv. Distance to default and parameters of distribution of V allow the computation of probability of default.

Formulas we use are

$$E = V\Phi(d_1) - D\exp(-T)\Phi(d_2) \tag{8.1}$$

From which

$$\sigma_e = \sigma_V \frac{V\Phi(d_1)}{E} \tag{8.2}$$

These two give V and σ_V.
Frequently used models for V are of the following class

$$\frac{dV(t)}{V(t)} = \mu(V, \sigma_V)dt + \sigma_V dz_t \tag{8.3}$$

8 Extreme value analysis

Suppose that losses $L_1, \ldots, L_N$ are independent with common d.f. $F_L(z) = \Pr\{L \leq z\}$. We are interested in the qualitative behavior of the distribution of the largest loss $M_N = \max(L_1, \ldots, L_N)$ for large values of N. The d.f. of M_N is simple, as

$$F_{M_N}(x) = \Pr\{M_N \leq x\} = \Pr\{L_1 \leq x, \ldots, L_N \leq x\} = [F_L(x)]^N.$$

From the practical viewpoint, this result is not very helpful as, for any x, $F_{M_N}(x)$ converges to either 0 or 1 as $N \to \infty$.

Suppose, however, that one can find a sequence of constants $\{(a_n, b_n)\}$ such that

$$\lim_{N\to\infty} \Pr\left\{\frac{M_N - a_N}{b_N} \leq z\right\} = H(z), \tag{8.80}$$

where H is some nondegenerate d.f. Then we may use H, called an *extreme value distribution*, to approximate the distribution of M_N in much the same way as the normal distribution provides an approximation to the sum of r.v.s. We shall not discuss in detail what are the conditions under which (8.80) holds, but just point out that they are not innocouos since they rule out respectable distributions such as the Poisson.

An important result, know as the *Fisher-Tippet Theorem* states that, if (8.80) holds for some H, then H can only be of three types:

i. **Fréchet type, or thick tails:**

$$\Phi_\alpha(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ \exp(-x^{-\alpha}), & \text{if } x > 0, \end{cases}$$

for some $\alpha > 0$.

ii. **Gumbel type, or normal tails:** $\Lambda(x) = \exp(-e^{-x})$.

iii. **Weibull case, or thin tails:**

$$\Psi_\alpha(x) = \begin{cases} \exp[-(-x)^\alpha], & \text{if } x \leq 0, \\ 0, & \text{if } x > 0, \end{cases}$$

for some $\alpha > 0$.

In credit-risk problems, i. and ii. are particularly relevant. It is worth pointing out, however, that the three distributions are closely related. In fact, if X is a r.v. with d.f. Φ_α, then $Y = \ln X^\alpha$ has d.f. Λ and $Z = -1/X$ has d.f. Ψ_α.

Fréchet type distributions include the Pareto, the Cauchy, the Burr and stable distributions with exponent less than 2. Gumbel type distributions include the exponential, the normal, the Erlang and the Weibull. Weibull type distributions include the uniform and the beta.

A second important result, known as the *Jenkinson-von Mises representation*, shows that the three possible types of extreme value distributions have the common one-parameter representation, called *generalized extreme value (GEV)* distribution

$$H_\gamma(x) = \begin{cases} \exp[-(1+\gamma x)_+^{-1/\gamma}], & \text{if } \gamma \neq 0, \\ \exp[-\exp(-x)], & \text{if } \gamma = 0. \end{cases}$$

with $u_+ = \max\{0, u\}$ denoting the positive part of u. The constant γ is known as the *extremal index*. The Fréchet case corresponds to $\gamma > 0$, the Gumbel case corresponds to $\gamma = 0$, and the Weibull case to $\gamma < 0$. Thus, the GEV provides a convenient unifying representation of the three types of extreme value distributions.

The support of H_γ is

$$\begin{array}{lll} x > -1/\gamma & \text{if} & \gamma > 0, \\ x \in \Re & \text{if} & \gamma = 0, \\ x < -1/\gamma & \text{if} & \gamma < 0. \end{array}$$

References

[1] Altman, E.I. (1989), "Measuring corporate bond mortality and performance", *Journal of Finance*, September, 909-922.

[2] Altman, E.I., G. Marco, and F. Varetto (1994), "Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience)", *Journal of Banking and Finance*, 18, 505–529.

[3] Beirlant, J., J.L. Teugels, and P. Vynckier (1996), *Practical Analysis of Extreme Values*, Leuven University Press, Leuven.

[4] Bessis, J. (1998), *Risk Management in Banking*, Wiley, Chichester.

[5] BIS (1995), *An Internal Model-Based Approach to Market Risk Capital Requirements*, Basle: Basle Committee on Banking Supervision, Bank of International Settlements.

[6] BIS (1996a), *Amendment to the Capital Accord to Incorporate Market Risk*, Basle: Basle Committee on Banking Supervision, Bank of International Settlements.

[7] BIS (1996b), *Overview of the Amendment to the Capital Accord to Incorporate Market Risk*, Basle: Basle Committee on Banking Supervision, Bank of International Settlements.

[8] Credit Suisse Financial Products (1997), *CrediRisk$^+$. A Credit Risk Management Framework*, London.

[9] Demirgüc-Kunt, A., (1992), "Deposit-Institution Failures: A Review of the Empirical Literature ", *Economic Review*, Federal Reserve Bank of Cleveland, pp. 2-18.

[10] Drudi, F., A. Generale, and G. Majnoni (1997), "Sensitivity of VaR measures to different risk models", Banca d'Italia, *Temi di Discussione* No. 317.

[11] Embrechts, P., C. Klüppelberg, and T. Mikosch (1997), *Modelling Extremal Events for Insurance and Finance*, Springer, Berlin.

[12] Feller W. (1971), *Introduction to Probability Theory and Its Applications*, Vol. 2 (2nd ed.) Wiley, New York.

[13] Gordy, M.B. (1998), "A comparative anatomy of credit risk models", Board of Governors of the Federal Reserve System, mimeo.

[14] Iben, T. and Litterman, R. (1989), "Corporate bond evaluation and the term structure of credit spreads", *Journal of Portfolio Management*, 52-64.

[15] Laviola, S., and M. Trapanese (1997), "Previsione delle insolvenze delle imprese e qualità del credito bancario: un'analisi statistica", Banca d'Italia, *Temi di Discussione* No. 318.

[16] Lawrence, E.L., Smith, S. and M. Rhoades (1992), "An analysis of default risk in mobile home credit", *Journal of Banking and Finance*, 299-312.

[17] Leland, H. (1994), "Risky Debt, Bond Covenants and Optimal Capital Structure", *Journal of Finance*, 49, 1213-1252.

[18] Merton, R. (1974), "On the pricing of corporate debit: the risk structure of interest rates", *Journal of Finance*, 29, 449–470.

[19] J.P. Morgan & Co. Incorporated (1997), *CreditMetrics*^{TM} - *Technical Document.*

[20] Parzen, E. (1979), "Nonparametric statistical data modeling", *Journal of the American Statistical Association*, 74, 105–131.

[21] Peracchi, F. (1995), *Econometria*, McGraw-Hill, Milan.

[22] Portnoy, S., and J. Jureckovà (1999), "On extreme regression quantiles", mimeo.

[23] Resti, A. (1999), "Misurare il rischio del portafoglio impieghi: introduzione ai modelli VaR per i crediti bancari", *Materiali* N. 25, Associazione Nazionale Banche Private, Milano.

[24] Smith, R. (1994), "Nonregular regression", *Biometrika*, 81, 173–183.

[25] Smith, L.D. and E. Lawrence, (1995), "Forecasting losses on a liquidating long-term loan portfolio", *Journal of Banking and Finance*, 959-95.