model risk is amplified by the cost to build it, to verify and maintain its effectiveness over time. If we try to avoid (even partially) these costs, this could reduce the model efficiency and accuracy.

Despite these problems, we have no real alternatives to using a firm's simulation model in specific conditions, such as when we have to analyze a start-up, and we have no means to observe historical data. Let's also think about special purpose entities, project companies, or companies that have merged recently, LBOs or other situations in which we have to assess plans but not facts. Moreover, in these transactions covenants and 'negative pledges' are very often contractually signed to control specific risky events, putting lenders in a better position to promptly act in deteriorating circumstances. These contractual clauses have to be modeled and contingently assessed to verify both when they are triggered and what their effectiveness is. The primary cash flow source for debt repayment stays in operational profits and, in case of difficulties, the only other source of funds is company assets; usually, these are no-recourse transactions and any guarantee is offered by third parties. These deals have to be evaluated only against future plans, with no past history backing-up lenders. Individual analysis is needed, and it is necessarily expensive. Therefore, these are often 'big ticket' transactions, to spread fixed costs on large amounts. Rating (and therefore default probability) is assigned using cash flow simulation models.

These models are often based on codified steps, producing inter-temporal specifications of future pro-forma financial reports, taking into consideration scenarios regarding:

- how much cash flows (a) will be generated by operations, (b) will be used for financial obligations and other investments, and (c) what are their determinants (demand, costs, technology, and other crucial hypotheses),
- complete future pro-forma specifications (at least ill the most advanced models), useful for also supporting more traditional analysis by ratios as well as for setting covenants and controls on specific balance sheet items.

To reach the probability of default, we can use either a scenario approach or a numerical simulation model. In the first case, we can apply probability to different (discrete) pre-defined scenarios. Rating will be determined through a weighted mathematical expectation on future outcomes; having a spectrum of future outcomes, defined by an associated probability of occurrence, we could also select a confidence level (e.g., 68% or 95%) to cautiously set our expectations. In the second case, we can use a large number of model iterations, which describe different scenarios: default and no-default (and also more diversified situations such as near-to-default, stressed and so forth) are determined and then the relative frequency of different stages is computed.

To give an example, Figure 4-7 depicts the architecture of a proprietary model developed for financial project applications, called SIMFLUX. The default criterion is defined in a Merton style approach; default occurs when the assets value falls below the 'debt barrier'. The model also proposes the market value of debt, showing all the intermediate stages (sharp reduction in debt value) in which repayment is challenging but still achievable (near-to-default stages). These situations are very important in financial projects because, very often, waivers are forced if things turn out badly, in order to minimize non-payment and default filing that would generate credit losses.

The model works as follows:

- the impact on project outcomes is measured, based on industrial valuations, sector perspectives and analysis, key success factors and so forth. Sensitivity to crucial macro-economic variables is then estimated. Correlation among the macro-economic risk factors is ascertained in order to find joint probabilities of potential future outcomes (scenario engine);
- given macroeconomic joint probabilities, random scenarios are simulated to generate revenues' volatility and its probability density function;
- applying the operational leverage, margin volatility is estimated as well. Then, the discount rate is calculated, with regards to the market risk premium and business volatility;
- applying the discount rate to cash flows, the firm's value is produced (from time to time for the first five years plus 'terminal value' beyond this horizon, using an asymptotic 'fading factor' for margins growth);
- Monte Carlo random simulations are then run, to generate the final expected spectrum of assets and debt value;
- then, default frequencies are counted, that is, the number of occurrences in which assets values are less than

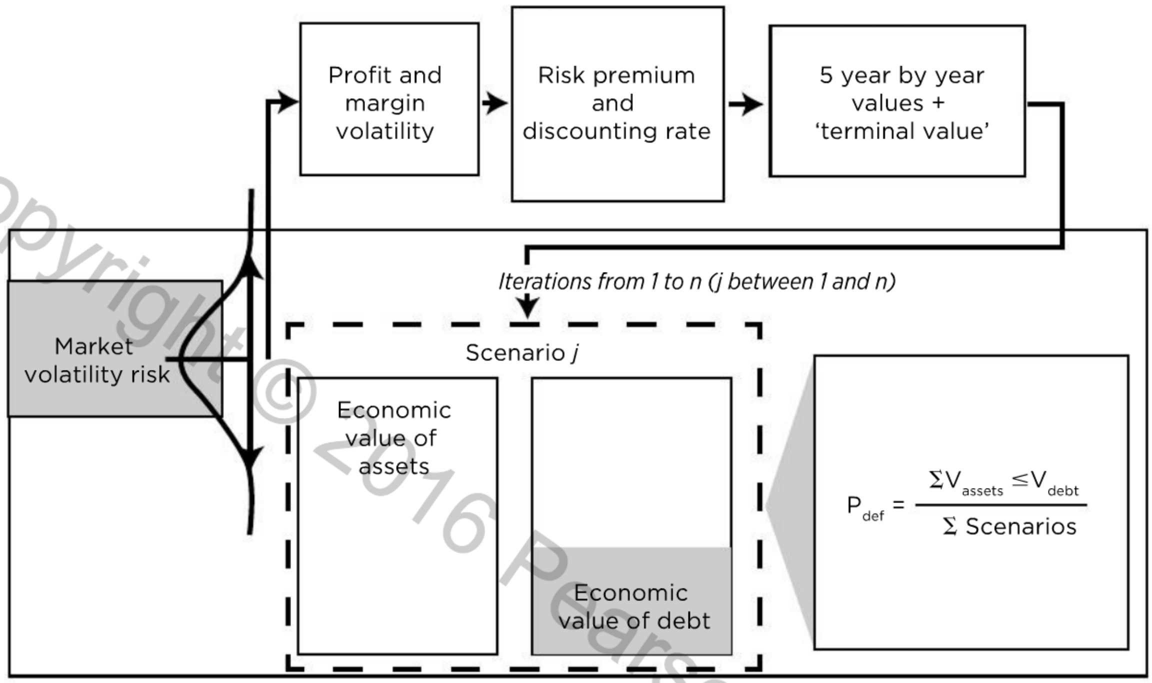


FIGURE 4-7 SIMFLUX cashflow simulation model architecture.

Source: Internally developed.

debt values. Consequently, a probability of default is determined;

- debt market values are also utilized by the model, plotted on a graph, to directly assess when there is a significant reduction in debt value, indicating difficulties and potential 'near-to-default' situations.

A Synthetic Vision of Quantitative-Based Statistical Models

Table 4-15 shows a summary valuation on quantitative statistical-based methods to ratings assignment, mapped against the three desirable features previously described.

Structural approaches are typically applied to listed companies due to the input data which is required. Variance reduction techniques are generally not seen as an alternative to regression or discriminant functions but rather as a complement of them: only cluster analysis can be considered as an alternative when top down approaches are preferred; this is the case when a limited range of data representing borrowers' characteristics is available. Cash flow analysis is used to rate companies whose track records are meaningless or non-existent. Discriminant and regression analyses are the principal techniques for bottom up statistical based rating models.

HEURISTIC AND NUMERICAL APPROACHES

In recent years, other techniques besides statistical analyses have been applied to default prediction; they are mostly driven by the application of artificial intelligence methods. These methods completely change the approach to traditional problem solving methods based on decision theory. There are two main approaches used in credit risk management:

1. 'Heuristic methods', which essentially mimic human decision making procedures, applying properly calibrated rules in order to achieve solutions in complex environments. New knowledge is generated on a trial by error basis, rather than by statistical modeling; efficiency and speed of calculation are critical. These methods are opposed to algorithms-based

TABLE 4-15 Overview of Quantitative-Based Statistical Ratings

Criteria	Structural Approach	Reduced Form Approaches			
	Option Approach Applied to Stock Listed Companies	Discriminant Analysis	Logistic Regression	Unsupervised Techniques*	Cash Flow Simulations
Measurability and verifiability	●	●	●	◕	◕
Objectivity and homogeneity	●	◕	◕	◑	◕
Specificity	◕	●	●	◔	◑

*Cluster analysis, principal components, factor analysis, canonical correlation.

approaches and are often known as 'expert systems' based on artificial intelligence techniques. The aim is to reproduce high frequency standardized decisions at the best level of quality and adopting low cost processes. Feedbacks are used to continuously train the heuristic system, which learns from errors and successes.

2. 'Numerical methods', whose objective is to reach optimal solutions adopting 'trained' algorithms to take decisions in highly complex environments characterized by inefficient, redundant, and fuzzy information. One example of these approaches is 'Neural networks': these are able to continuously auto-update themselves in order to adjust to environmental modifications. Efficiency criteria are externally given or endogenously defined by the system itself.

Expert Systems

Essentially, expert systems are software solutions that attempt to provide an answer to problems where human experts would need to be consulted. Expert systems are traditional applications of artificial intelligence. A wide variety of methods can be used to simulate the performance of an expert. Elements common to most or all expert systems are:

- the creation of a knowledge base (in other words, they are *knowledge-based systems),*
- the process of gathering knowledge and codifying it according to some frameworks (this is called knowledge engineering).

Hence, expert systems' typical components are:

1. the knowledge base,
2. the working memory,
3. the inferential engine,
4. the user's interface and communication.

The knowledge base is also known as 'long term memory' because it is the set of rules used for decisions making processes. Its structure is very similar to a database containing facts, measures, and rules, which are useful to tackle a new decision using previous (successful) experiences. The typical formalization is based on 'production rules', that is, 'if/then' hierarchical items, often integrated by probabilities p and utility u. These rules create a decision making environment that emulates human problem solving approaches. The speed of computers allows the application of these decision processes with high frequency in various contexts and circumstances, in a reliable and cost effective way.

The production of these rules is developed by specialists known as 'knowledge engineers'. Their role is to formalize the decision process, encapsulating the decision making logics and information needs taken from practitioners who are experts in the field, and finally combining different rules in layers of inter-depending steps or in decisional trees.

The 'working memory' (also known as short term memory) contains information on the problem to be solved and is, therefore, the virtual space in which rules are combined and where final solutions are produced. In recent years, information systems are no longer a constraint to the application of these techniques; computers' data storage capacity has increased to a point where it is possible to run certain types of simple expert systems even on personal computers.

The inferential engine, at the same time, is the heart and the nervous network of an expert system. An understanding of the 'inference rules' is important to comprehend how expert systems work and what they are useful for. Rules give expert systems the ability to find solutions to diagnostic and prescriptive problems. An expert system's rule-base is made up of many inference rules. They are entered into the knowledge base as separate rules and the inference engine uses them together to draw conclusions. As each rule is a unit, rules may be deleted or added without affecting other rules. One advantage of inference rules over traditional models is that inference rules more closely resemble human behavior. Thus, when a conclusion is drawn, it is possible to understand how this conclusion was reached. Furthermore, because the expert system uses information in a similar manner to experts, it may be easier to find out the needed information from banks' files. Rules can also incorporate probability of events and the gain/cost of them (utility).

The inferential engine may use two different approaches, backward chaining and forward chaining respectively:

- 'Forward chaining' starts with available data. Inference rules are used until a desired goal is reached. An inference engine searches through the inference rules until it finds a solution that is pre-defined as correct; the path, once recognized as successful, is then applied to data. Because available data determine which inference rules are used, this method is also known as data driven.
- 'Backward chaining' starts with a list of goals. Then, working backwards, the system tries to find the path which allows it to achieve any of these goals. An inferential engine using backward chaining would search through the rules until it finds the rule which best matches a desired goal. Because the list of goals determines which rules are selected and used, this method is also known as goal driven.

Using chaining methods, expert systems can also explore new paths in order to optimize target solutions over time.

Expert systems may also include fuzzy logic applications. Fuzzy logic has been applied to many fields, from control theory to artificial intelligence. In default risk analysis, many rules are simply 'rule-of-thumb' that have been derived from experts' own feelings; often, thresholds are set for ratios but, because of the complexity of real world, they can result to be both sharp and severe in many circumstances. Fuzzy logic is derived from 'fuzzy set theory', which is able to deal with approximate rather than precise reasoning. Fuzzy logic variables are not constrained to the two classic extremes of black and white logic (zero and one), but rather they may assume any value between the extremes. When there are several rules, the set thresholds can hide incoherencies or contradictions because of overlapping areas of uncertainty and logical mutual exclusions. Instead, adopting a more flexible approach, many clues can be integrated, reaching a solution that converges to a sounder final judgment. For example:

- if interest coverage ratio (EBIT divided by interest paid) is less than 1.5, the company is considered as risky,
- if ROS (EBIT divided by revenues) is more than 20%, the company is considered to be safe.

The two rules can be combined together. Only when both are valid, that is to say ROS is lower (higher) than 20% and interest coverage is lower (higher) than 1.5, we can reach a dichotomous risky/safe solution. In all other cases, we are uncertain.

When using the fuzzy logic approach, the 'low interest coverage rule' may assume different levels depending on ROS. So, when a highly profitable company is considered, less safety in interest coverage can be accepted (for instance, 1.2). Therefore, fuzzy logic widens the spectrum of rules that expert systems can use, allowing them to approximate human decisional processes even better.

Expert systems were created to substitute human-based processes by applying mechanical and automatic tools. When knowledge is well consolidated and stabilized, characterized by frequent (complex and recursive) calculations and associated with well established decision rules, then expert systems are doing their best in exploring all possible solutions (may be millions) and in finding out the best one. Over time, their knowledge base has extended to also include ordinal and qualitative information as well as combinations of statistical models, numerical methods, complex algorithms, and logic/hierarchical patterns of many interconnected submodels. Nowadays, expert systems are more than just a way to solve problems or to model some real world phenomena; they are software that connects many subprocesses and procedures, each optimized in relation to its goals using different rules. Occasionally, expert systems are also used when there are completely new conditions unknown to the human experience (new products, new markets, new procedures, and so forth). In these cases, as there is no expertise, we need to explore what can be achieved by applying rules derived from other contexts and by following a heuristic approach.

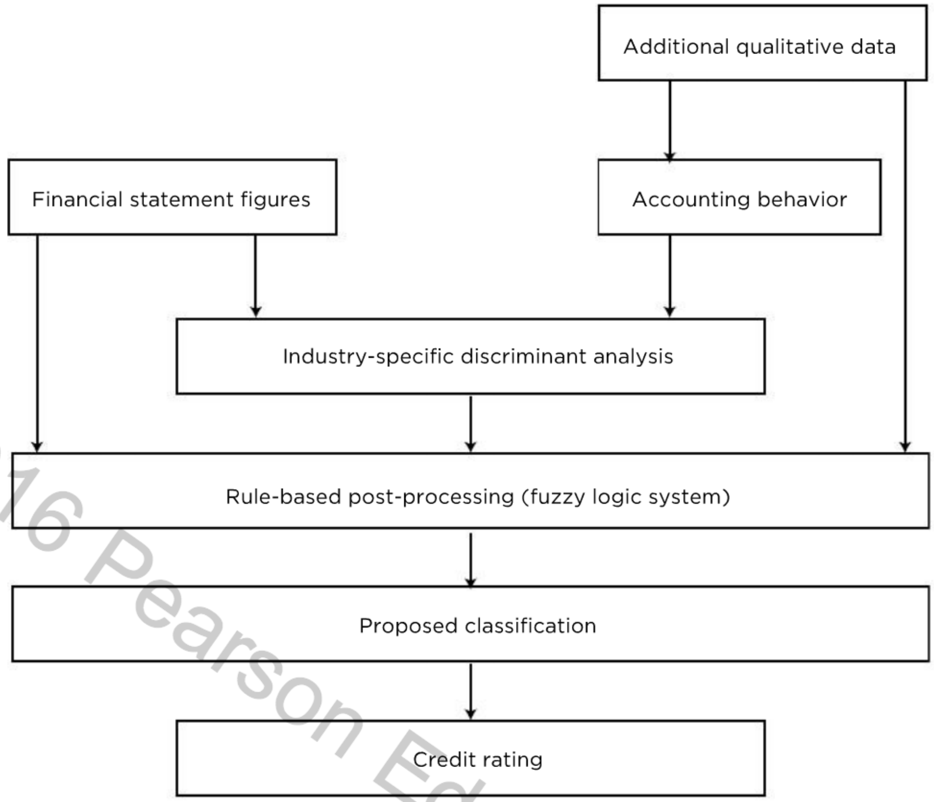


FIGURE 4-8 Expert system in Bundesbank's credit quality valuation model.

In the credit risk management field, an expert system based on fuzzy logic used by the German Bundesbank since 1999 (Blochwitz and Eigermann, 2000) is worth noting. It was used in combination with discriminant analysis to investigate companies that were classified by the discriminant model in the so-called 'gray area' (uncertain attribution to defaulting/performing classes). The application of the expert system raised the accuracy from 18.7% of misclassified cases by discriminant function to an error rate of only 16% for the overall model (Figure 4-8).

In the early 1990s, SanPaoloIMI also built an expert system for credit valuation purposes, based on approximately 600 formal rules, 50 financial ratios, and three areas of analysis. According to a blind test, the system proved to be very efficient and guaranteed homogeneity and a good quality of results. Students who were studying economics (skilled in credit and finance) and students studying engineering (completely unskilled in credit and finance) were asked to separately apply the expert system to about 100 credit dossiers without any interaction. The final result was remarkable. The accuracy in data loading and in output produced, the technical comments on borrowers' conditions, and the time employed, were the same for both groups of skilled and unskilled people. In addition, the borrowers' defaulting or performing classifications were identical because they were strictly depending on the model itself. These are, at the same time, very clearly, the limits and opportunities of expert systems.

Decision Support Systems (DSSs) are a subset of expert systems. These are models applied to some phases of the human decision process, which mostly require cumbersome and complex calculations. DSSs have had a certain success in the past, also as stand-alone solutions, when computer power was quickly increasing. Today, many DSSs applications are part of complex procedures, supporting credit approval processes and commercial relationships.

Neural Networks

Artificial neural networks originate from biological studies and aim to simulate the behavior of the human brain, or at least a part of the biological nervous system (Arbib, 1995; Steeb, 2008). They comprise interconnecting artificial neurons, which are software programs intended to mimic the properties of biological neurons.

Artificial neurons are hierarchical 'nodes' (or steps) connected in a network by mathematical models that are able to exploit connections by operating a mathematical transformation of information at each node, often adopting a fuzzy logic approach. In Figure 4-9, the network is reduced to its core, that is, three layers. The first is delegated to handle inputs, pre-filtering information, stimuli, and signals. The second (hidden) is devoted to computing relationships and aggregations; in more complex neural networks this component could have many layers. The third is designated to generate outputs and to manage the users' interface, delivering results to the following processes (human or still automatic).

Thanks to the (often complex) network of many nodes, the system is able to fit into many different cases and can also describe nonlinear relationships in a very flexible way. After the 'initial training', the system can also improve its adaptability, learning from its successes and failures over time.

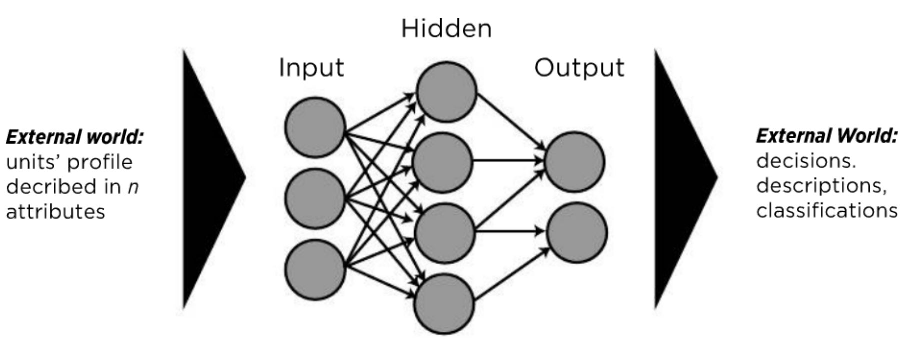


FIGURE 4-9 Frame of a neural network.

To better clarify these concepts, compare neural networks with traditional statistical analyses such as regression analysis. In a regression model, data are fitted through a specified relationship, usually linear. The model is made up by one or more equations, in which each of the inputs x_j is multiplied by a weight w_j. Consequently, the sum of all such products and of a constant α gives an estimate of the output. This formulation is stable over time and can only be changed by an external decision of the model builder.

In neural networks, the input data x_j is again multiplied by weights (also defined as the 'intensity' or 'potential' of the specific neuron), but the sum of all these products is influenced by:

- the argument of a flexible mathematical function (e.g., hyperbolic tangent or logistic function),
- the specific calculation path that involves some nodes, while ignoring others.

The network calculates the signals gathered and applies a defined weight to inputs at each node. If a specific threshold is overcome, the neuron is 'active' and generates an input to other nodes, otherwise it is ignored. Neurons can interact with strong or weak connections. These connections are based on weights and on paths that inputs have to go through before arriving to the specific neuron. Some paths are privileged; however neurons never sleep: inputs always go through the entire network. If new information arrives, the network can search for new solutions testing other paths, and thus activating certain neurons while switching off others. Some paths could automatically substitute others because of the change in input intensity or in inputs profile. Often, we are not able to perceive these new paths because the network is always 'awake', immediately catching news and continuously changing neural distribution of stimuli and reactions. Therefore, the output y is a nonlinear function of x_j. So, the neural network method is able to capture nonlinear relationships.

Neural networks could have thousands of nodes and, therefore, tens of thousands of potential connections. This gives great flexibility to the whole process to tackle very complex, interdependent, no linear, and recursive problems.

The most commonly used structure is the 'hierarchically dependent neural network'. Each neuron is connected with previous nodes and delivers inputs to the following node, with no return and feedbacks, in a continuous and ordered flow. The final result is, therefore, the nonlinear weighted sum of inputs and defined as:

$$f(x) = k\left(\sum_i w_i g_i(x)\right)$$

where k is a pre-defined function; for instance, the logistic one. The ith neuron gathers stimuli from j previous neurons. Based on weights, the 'potential', v_i is calculated in a way depicted in Figure 4-10.

The potential is not comparable among neurons. A conversion is needed in order to compare them; the logistic conversion indicated in Figure 4-10 sets the output value between 0 and 1. When there is only one hidden layer, the neural network behaves like a traditional statistical logistic function. Unless very complex problems are being dealt with, one or two layers are enough to solve most issues, as also proven by mathematical demonstrations.

The final result of the neural network depends more on training than on the complexity of the structure.

Now, let's come to the most interesting feature of a neural network, the ability to continuously learn by experience (neural networks are part of 'learning machines'). There are different learning methods and many algorithms for training neural networks. Most of them can be viewed as a straightforward application of the optimization theory and statistical estimation.

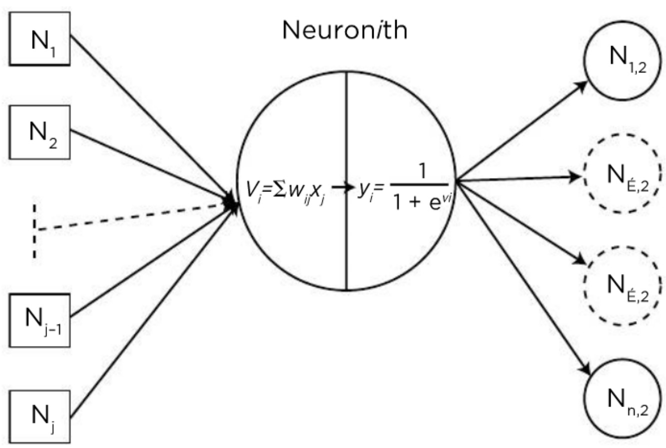


FIGURE 4-10 Communications among artificial neurons.

In the field of credit risk, the most applied method is 'supervised learning', in which the training set is given and the neural network learns how to reach a successful result by finding the nodes' structure and the optimal path to reach the best final result. This also implies that a cost function is set in order to define the utility of each outcome. In the case of default risk model building, the training set is formed by borrowers' characteristics and the cost function reflects misclassification costs. A backpropagation learning engine may be launched to train the neural network. After much iteration, a solution that minimizes the classification errors is reached by changing the weights and connections at different nodes. If the training process is successful, the neural network learns the connections among inputs and outputs, and can be used to make previsions for new borrowers that were not present in the training set. Accuracy tests are to be performed to gauge if the network is really able to solve problems in out-of-sample populations, with an adequate level of generality.

The new generations of neural networks are more and more entwined with statistical models and numerical methods. Neural networks are (apparently) easy to use and generate a workable solution fairly quickly. This could set a mental trap: complex problems remain complex even if a machine generates adequate results; competences in statistics and a good control of the information set are unavoidable.

The main limit of neural networks is that we have to accept results from a 'black box'. We cannot examine step by step how results are obtained. Results have to be accepted as they are. In other words, we are not able to explain why we arrive at a given result. The only possibility is to prepare various sets of data, well characterized with some distinguishing profiles, then submit them to the neural network to reach results. In this case, by having outputs corresponding to homogenous inputs and using the system theory, we can deduce which the crucial variables are and their relative weights.

Much like any other model, neural networks are very sensitive to input quality. So, training datasets have to be carefully selected in order to avoid training the model to learn from outliers instead of normal cases. Another limit is related to the use of qualitative variables. Neural networks are more suited to work with continuous quantitative variables. If we use qualitative variables, it is advisable to avoid dichotomous categorical variables, preferring multiple ranked modalities.

There are no robust scientific ways to assess if a neural network is optimally estimated (after the training process). The judgment on the quality of the final neural network structure is mainly a matter of experience, largely depending on the choices of knowledge engineers made during model building stages.

The major danger in estimating neural networks is the risk of 'over-fitting'. This is a sort of network over-specialization in interpreting the training sample; the network becomes completely dependent on the specific training set. A network that over-fits a sample is incapable of producing satisfactory results when applied to other borrowers, sectors, geographical areas or economic cycle stages. Unfortunately, there are no tests or techniques to gauge if the solution is actually over-fitting or not. The only way out is to use practical solutions. On one hand, the neural network has to be applied to out-of-sample, out-of-time and out-of-universe datasets to verify if there are significant falls in statistical performances. On the other hand, the neural network has to be continuously challenged, very often re-launching the training process, changing the training set, and avoiding specialization in only one or few samples.

In reality, neural networks are mainly used where decisions are taken in a fuzzy environment, when data are rough, sometimes partially missed, unreliable, or mistaken. Another elective realm of application lies where dominant approaches are not provided, because of the complexity, novelty or rapid changes in external conditions. Neural networks are, for instance, nowadays used in negotiation platforms. They react very quickly to changing market conditions. Their added value clearly stays in a prompt adaptation to structural changes.

Comparison of Heuristic and Numerical Approaches

Expert systems offer advantages when human experts' experience is clear, known, and well dominated. This enables knowledge engineers to formalize rules and build effective systems. Expert systems are ideal for the following reasons:

- to give order and structure to real life procedures, which allow decision making processes to be replicated with high frequency and robust quality;
- to connect different steps of decision making processes to one another, linking statistical and inferential engines, procedures, classifications, and human involvement together, sometimes reaching the extreme of producing outputs in natural language.

From our perspective (rating assignment), expert systems have the distinct advantage of giving order, objectivity, and discipline to the rating process; they are desirable features in some circumstances but they are not decisive. In reality, rating assessment depends more on models that are implanted inside the expert system, models that are very often derived from other methods (statistical, numerical), with their own strength and weaknesses. As a result, expert systems organize knowledge and processes but they do not produce new knowledge because they are not models or inferential methods.

Numerical algorithms such as neural networks have completely different profiles. Some applications perform quite satisfactorily and count many real life applications. Their limits are completely different and are mainly attributable to the fact that they are not statistical models and do not produce a probability of default. The output is a classification, sometimes with a very low granularity (four classes, for instance, such as very good, pass, verify, reject). To extract a probability, we need to associate a measure of default frequency obtained from historical data to each class. Only in some advanced applications, does the use of models implanted in the inferential engine (i.e., a logistic function) generate a probability of default. This disadvantage, added to the 'black box nature' of the method, limits the diffusion of neural networks out of consumer credit or personal loan segments.

However, it is important to mention their potential application in early warning activities and in credit quality monitoring. These activities are applied to data generated by very different files (in terms of structure, use, and scope) that need to be monitored frequently (i.e., daily or even intraday for internal behavioral data). Many databases often derived from production processes frequently change their structure. Neural networks are very suitable

TABLE 4-16 An Overview of Heuristic and Numerical Based Ratings

Criteria	Heuristic Approach Expert Systems/Decision Support System	Numerical Approach Neural Networks
Measurability and verifiability	◑	◕
Objectivity and homogeneity	●	●
Specificity	◑	◕

in going through a massive quantity of data, changing rapidly when important discontinuities occur, and quickly developing new rules when a changing pattern of success/failure is detected.

Finally, it should be noted that some people classify neural networks among statistical methods and not among numerical methods, because some nodes can be based on statistical models (for instance, ONB and FMA, 2004).

Table 4-16 offers the usual final graphic summary:

INVOLVING QUALITATIVE INFORMATION

Statistical methods are well suited to manage quantitative data. However, useful information for assessing probability of default is not only quantitative. Other types of information are also highly relevant such as: sectors competitive forces characteristics, firms' competitive strengths and weaknesses, management quality, cohesion and stability of entrepreneurs and owners, managerial reputation, succession plans in case of managerial/entrepreneurial resignation or turnaround, strategic continuity, regulations on product quality, consumers' protection rules and risks, industrial cost structures, unionization, non-quantitative financial risk profiles (existence of contingent plans on liquidity and debt repayment, dependency from the financial group's support, strategy on financing growth, restructuring and repositioning) and so forth; these usually have a large role in judgment-based approaches to credit approval and can be classified in three large categories:

1. efficiency and effectiveness of internal processes (production, administration, marketing, post-marketing, and control);
2. investment, technology, and innovation;
3. human resource management, talent valorization, key resources retention, and motivation.

In more depth:

- domestic market, product/service range, firm's and entrepreneurial history and perspectives;
- main suppliers and customers, both in terms of quality and concentration;
- commercial network, marketing organization, presence in regions and countries, potential diversification;
- entrepreneurial and managerial quality, experience, competence;
- group organization, like group structure, objectives and nature of different group's entities, main interests, diversification in non-core activities, if any;

- investments in progress, their final foreseeable results in maintaining/re-launching competitive advantages, plans, and programs for future competitiveness;
- internal organization and functions, resources allocation, managerial power, internal composition among different branches (administration, production, marketing, R&D and so forth);
- past use of extraordinary measures, like government support, public wage integration, foreclosures, payments delays, credit losses and so forth;
- financial relationships (how many banks are involved, quality of relationships, transparency, fairness and correctness, and so on);
- use of innovative technologies in payment systems, integration with administration, accounting, and managerial information systems;
- quality of financial reports, accounting systems, auditors, span of information and transparency, internal controls, managerial support, internal reporting and so on.

Presently, new items have become of particular importance: environmental compliance and conformity, social responsibility, corporate governance, internal checks and balances, minorities protection, hidden liabilities like pension funds integration, stock options and so on.

A recent summing up of usual qualitative information conducted in the Sanpaolo Group in 2006 collected more than 250 questions used for credit approval processes, extracted from the available documents and derived from industrial and financial economy, theories of industrial competition and credit analysis practices. A summary is given in Table 4-17.

Qualitative variables are potentially numerous and, consequently, some ordering criterion is needed to avoid complex calculations and information overlapping. Moreover, forms to be filled in soon become very complex and difficult to be understood by analysts. A first recommendation is to only gather qualitative information that is not collectable in quantitative terms. For instance, growth and financial structure information can be extracted from balance sheets.

A second recommendation regards how to manage qualitative information in quantitative models. A preliminary distinction is needed between different categorical types of information:

- nominal information, such as regions of incorporation;
- binary information (yes/no, presence/absence of an attribute);
- ordinal classification, with some graduation (linear or nonlinear) in the levels (for instance, very low/low/medium/high/very high).

Binary indicators can be transformed in 0/1 *'dummy variables'.* Also, ordinal indicators can be transformed into numbers and weights can be assigned to different modalities (the choice of weights is, however, debatable).

When collecting data, it is preferable to structure the information in closed form if we want to use it in quantitative models. This means forcing loan officers to select some pre-defined answers.

Binary variables are difficult to manage in statistical models because of their non-normal distribution. Where possible, a multistage answer is preferable, instead of yes/no. Weights can be set using optimization techniques, like *'bootstrap'* or a preliminary test on different solutions to select the most suited one.

Nowadays, however, the major problem in using qualitative information lies in the lack of historical datasets. The credit dossier is often based on literary presentations without a structured compulsory basic scheme. Launching an extraordinary survey in order to collect missing information has generally proven to be:

- A very expensive solution. There are thousands of dossiers and it takes a long time for analysts to go through them. Final results may be inaccurate because they are generated under pressure.
- A questionable approach for non-performing dossiers. Loan and credit officers that are asked to give judgments on the situation as it was before the default are tempted to skip on weaknesses and to hide true motivations of their (ex post proved wrong) judgments.

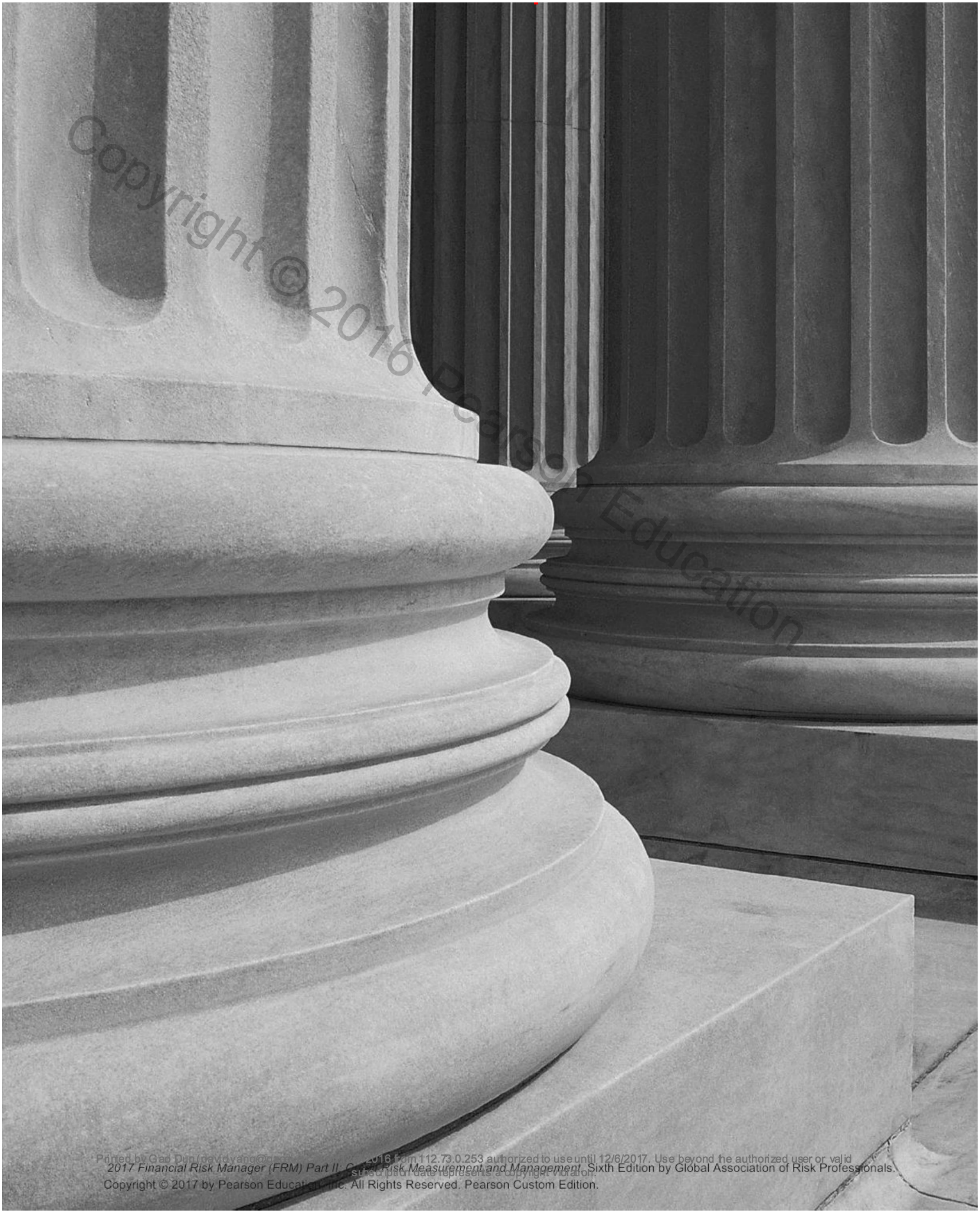
TABLE 4-17 Example of Qualitative Items in Credit Analysis Questionnaires

• Corporate structure —date of incorporation of the company (or of a significant merger and/or acquisition) —group members, intensity of relationship with the parent/subsidiary
• Information on the company's business —markets in which the company operates and their position in the 'business life cycle' (introduction, growth, consolidation, decline) —positions with competitors and competitive strength —nature of competitive advantage (cost, differentiation/distinctiveness of products, quality/innovation/technology, dominant/defendable) —years the company operates in the actual core business —growth forecast —quality of the references in the marketplace
• Strategy —strategic plans —business plan —in case a business plan has been developed, the stage of strategy implementation —proportion of assets/investments not strategically linked to the company's business —extraordinary transactions (revaluations, mergers, divisions, transfers of business divisions, demerger of business) and their objective
• Quality of management —degree of involvement in the ownership and management of the company —the overall assessment of management's knowledge, experience, qualifications and competence (in relation to competitors) —if the company's future is tied to key figures —presence of a dominant entrepreneur/investor (or a coordinated and cohesive group of investors) that influence strategies and company's critical choices
• Other risks —risks related to commercial activity —geographical focus (the local/regional, domestic, within Europe, OECD and non-OECD/emerging markets) —level of business diversification (a single product/service, more products, services, markets) —liquidity of inventories —quality of client base —share of total revenues generated by the first three/five customers of the company —exclusivity or prevalence with some company's suppliers —legal and/or environmental risks —reserves against professional risks, board members responsibilities, auditors (or equivalent insurance)
• Sustainability of financial position —reimbursements within the next 12 months, 18 months, 3 years, and concentration of any significant debt maturities —off-balance-sheet positions and motivations (coverage, management, speculation, other) —sustainability of critical deadlines with internal/external sources and contingency plans —liquidity risk, potential loss in receivables of one or more major customers (potential need to accelerate the payment of the most important suppliers)
• Quality of information provided by the company to the bank, timing in the documentation released and general quality of relationships —availability of plausible financial projections —information submitted on company's results and projections —considerations released by auditors on the quality of budgetary information —relationship vintage, past litigation, type of relation (privileged/strategic or tactical/opportunistic) —managerial attention —negative signals in the relationship history

There is no easy way to overcome these problems. A possible way is to prepare a two-stage process:

- The first stage is devoted to building a quantitative model accompanied by the launch of a systematic qualitative data collection on new dossiers. This qualitative information can immediately be used in overriding quantitative model results through a formal or informal procedure.
- The second stage is to build a new model including the new qualitative information gathered once the first stage has produced enough information (presumably after at least three years), trying to find the most meaningful data and possibly re-engineering the data collection form if needed.

Note that qualitative information change weight and meaningfulness over time. At the end of the 1980s, for instance, one of the most discriminant variables was to operate or not on international markets. After the globalization, this feature is less important; instead, technology, marketing skills, brands, quality, and management competences have become crucial. Therefore, today, a well structured and reliable qualitative dataset is an important competitive hedge for banks, an important component to build powerful credit models, and a driver of banks' long term value creation.



Credit Risks and Credit Derivatives

5

Learning Objectives

After completing this reading you should be able to:

- Using the Merton model, calculate the value of a firm's debt and equity and the volatility of firm value.
- Explain the relationship between credit spreads, time to maturity, and interest rates.
- Explain the differences between valuing senior and subordinated debt using a contingent claim approach.
- Explain, from a contingent claim perspective, the impact of stochastic interest rates on the valuation of risky bonds, equity, and the risk of default.
- Compare and contrast different approaches to credit risk modeling, such as those related to the Merton model, CreditRisk+, CreditMetrics, and the KMV model.
- Assess the credit risks of derivatives.
- Describe a credit derivative, credit default swap, and total return swap.
- Explain how to account for credit risk exposure in valuing a swap.

Excerpt is Chapter 18 of Risk Management and Derivatives, *by René Stulz.*

Consider Credit Bank Corp. It makes loans to corporations. For each loan, there is some risk that the borrower will default, in which case Credit will not receive all the payments the borrower promised to make. Credit has to understand the risk of the individual loans it makes, but it must also be able to quantify the overall risk of its loan portfolio. Credit has a high franchise value and wants to protect that franchise value by making sure that the risk of default on its loans does not make its probability of financial distress too high. Though Credit knows how to compute VaR for its trading portfolio, it cannot use these techniques directly to compute the risk of its loan portfolio. Loans are like bonds—Credit never receives more from a borrower than the amounts the borrower promised to pay. Consequently, the distribution of the payments received by borrowers cannot be lognormal.

To manage the risk of its loans, Credit must know how to quantify the risk of default and of the losses it makes in the event of default both for individual loans and for its portfolio of loans. At the end of this chapter, you will know the techniques that Credit can use for this task. We will see that the Black-Scholes formula is useful to understand the risks of individual loans. Recently, a number of firms have developed models to analyze the risks of portfolios of loans and bonds. For example, J.P. Morgan has developed CreditMetrics™ along the lines of its product RiskMetrics™. We discuss this model in some detail.

A credit risk is the risk that someone who owes money might fail to make promised payments. Credit risks play two important roles in risk management. First, credit risks represent part of the risks a firm tries to manage in a risk management program. If a firm wants to avoid lower tail outcomes in its income, it must carefully evaluate the riskiness of the debt claims it holds against third parties and determine whether it can hedge these claims and how. Second, the firm holds positions in derivatives for the express purpose of risk management. The counterparties on these derivatives can default, in which case the firm does not get the payoffs it expects on its derivatives. A firm taking a position in a derivative must therefore evaluate the riskiness of the counterparty in the position and be able to assess how the riskiness of the counterparty affects the value of its derivatives positions.

Credit derivatives are one of the newest and most dynamic growth areas in the derivatives industry. At the end of 2000, the total notional amount of credit derivatives was estimated to be $810 billion; it was only $180 billion two years before. Credit derivatives have payoffs that depend on the realization of credit risks. For example, a credit derivative could promise to pay some amount if Citibank defaults and nothing otherwise; or a credit derivative could pay the holder of Citibank debt the shortfall that occurs if Citibank defaults on its debt. Thus firms can use credit derivatives to hedge credit risks.

CREDIT RISKS AS OPTIONS

Following Black and Scholes (1973), option pricing theory has been used to evaluate default risky debt in many different situations. The basic model to value risky debt using option pricing theory is the Merton (1974) model. To understand this approach, consider a levered firm that has only one debt issue and pays no dividends. Financial markets are assumed to be perfect. There are no taxes and no bankruptcy costs, and contracts can be enforced costlessly. Only debt holders and equity holders have claims against the firm and the value of the firm is equal to the sum of the value of debt and the value of equity. The debt has no coupons and matures at T.

At date T, the firm has to pay the principal amount of the debt, F. If the firm cannot pay the principal amount at T, it is bankrupt, equity has no value, and the firm belongs to the debt holders. If the firm can pay the principal at T, any dollar of firm value in excess of the principal belongs to the equity holders.

Suppose the firm has issued debt that requires it to make a payment of $100 million to debt holders at maturity and that the firm has no other creditors. If the total value of the firm at maturity is $120 million, the debt holders receive their promised payment, and the equity holders have $20 million. If the total value of the firm at maturity is $80 million, the equity holders receive nothing and the debt holders receive $80 million.

Since the equity holders receive something only if firm value exceeds the face value of the debt, they receive $V_T - F$ if that amount is positive and zero otherwise. This is equivalent to the payoff of a call option on the value of the firm. Let V_T be the value of the firm and S_T be the value of equity at date T. We have at date T:

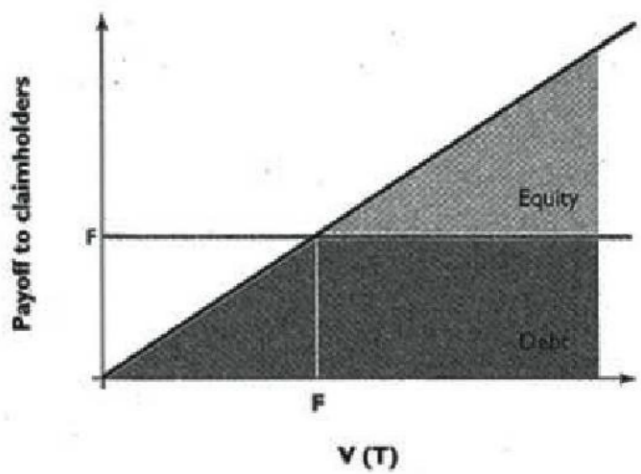
$$S_T = \text{Max}(V_T - F, 0) \tag{5.1}$$

To see that this works for our example, note that when firm value is $120 million, we have S_T equal to Max($120M − $100M, 0), or $20 million, and when firm value is $80 million, we have S_T equal to Max($80M − $100M, 0), or $0.

Figure 5-1 graphs the payoff of the debt and of the equity as a function of the value of the firm. If the debt were riskless, its payoff would be the same for any value of the firm and would be equal to F. Since the debt is risky, when the value of the firm falls below F, the debt holders receive less than F by an amount equal to $F - V_T$. The amount $F - V_T$ paid if V_T is smaller than F, $\text{Max}(F - V_T, 0)$, corresponds to the payoff of a put option on V_T with exercise price F. We can therefore think of the debt as paying F for sure minus the payoff of a put option on the firm with exercise price F:

$$D_T = F - \text{Max}(F - V_T, 0) \tag{5.2}$$

where D_T is the value of the debt at date T. Equation (5.2) therefore tells us that the payoff of risky debt is equal to the payoff of a long position in a risk-free zero-coupon bond with face value F and a short position on a put option on firm value with exercise price F. This means that holders of risky debt effectively buy risk-free debt but write a put option on the value of the firm with exercise price equal to the face value of the debt. Alternatively, we can say that debt holders receive the value of the firm, V_T, minus the value of equity, S_T. Since the payoff of equity is the payoff of a call option, the payoff of debt is the value of the firm minus the payoff of a call option with exercise price equal to the principal amount of the debt.



F is the debt principal amount and V(T) is the value of the firm at date T.

FIGURE 5-1 Debt and equity payoffs when debt is risky.

To price the equity and the debt using the Black-Scholes formula for the pricing of a European call option, we require that the value of the firm follow a log-normal distribution with a constant volatility s, the interest rate r be constant, trading take place continuously, and financial markets be perfect. We do not require that there is a security that trades continuously with value V. All we need is a portfolio strategy such that the portfolio has the same value as the firm at any particular time. We use this portfolio to hedge options on firm value, so that we can price such options by arbitrage. We can write the value of equity as S(V, F, T, *t*) and use the formula to price a call option to obtain:

Merton's Formula for the Value of Equity

Let S(V, F, T, *t*) be the value of equity at date *t*, V the value of the firm, F the face value of the firm's only zero-coupon debt maturing at T, σ the volatility of the value of the firm, $P_t(T)$ the price at *t* of a zero-coupon bond that pays $1 at T, and N(*d*) the cumulative distribution function evaluated at *d*. With this notation, the value of equity is:

$$S(V, F, T, t) = VN(d) - P_t(T)FN\left(d - \sigma\sqrt{T-t}\right)$$
$$d = \frac{\ln(V/P_t(T)F)}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t} \tag{5.3}$$

When V is $120 million, F is $100 million, T is equal to $t + 5$, $P_t(T)$ is $0.6065, and σ is 20 percent, the value of equity is $60.385 million. From our understanding of the determinants of the value of a call option, we know that equity increases in value when the value of the firm increases, when firm volatility increases, when time to maturity increases, when the interest rate increases, and when the face value amount of the debt falls.

Debt can be priced in two different ways. First, we can use the fact that the payoff of risky debt is equal to the payoff of risk-free debt minus the payoff of a put option on the firm with exercise price equal to the face value of the debt:

$$D(V, F, T, t) = P_t(T)F - p(V, F, T, t) \tag{5.4}$$

where *p*(V, F, T, *t*) is the price of a put with exercise price F on firm value V. F is $100 million and $P_t(T)$ is $0.6065, so $P_t(T)F$ is $60.65 million. A put on the value of the firm with exercise price of $100 million is worth $1.035 million. The value of the debt is therefore $60.65M − $1.035M, or $59.615 million.

The second approach to value the debt involves subtracting the value of equity from the value of the firm:

$$D(V, F, T, t) = V - S(V, F, T, t) \quad \textbf{(5.5)}$$

We subtract $60.385 million from $120 million, which gives us $59.615 million.

The value of the debt is, for a given value of the firm, a decreasing function of the value of equity, which is the value of a call option on the value of the firm. Everything else equal, therefore, the value of the debt falls if the volatility of the firm increases, if the interest rate rises, if the principal amount of the debt falls, and if the debt's time to maturity lengthens.

To understand the effect of the value of the firm on the value of the debt, note that a $1 increase in the value of the firm affects the right-hand side of Equation (5.5) as follows: It increases V by $1 and the call option by $\$\delta$, where δ is the call option delta. $1 - \delta$ is positive, so that the impact of a $1 increase in the value of the firm on the value of the debt is positive and equal to $\$1 - \$\delta$. δ increases as the call option gets more in the money. Here, the call option corresponding to equity is more in the money as the value of the firm increases, so that the impact of an increase in firm value on debt value falls as the value of the firm increases.

Investors pay a lot of attention to credit spreads. The credit spread is the difference between the yield on the risky debt and the yield on risk-free debt of same maturity. If corporate bonds with an A rating have a yield of 8 percent while T-bonds of the same maturity have a yield of 7 percent, the credit spread for A-rated debt is 1 percentage point. An investor can look at credit spreads for different ratings to see how the yields differ across ratings classes. An explicit formula for the credit spread is:

$$\text{Credit spread} = -\left(\frac{1}{T-t}\right)\ln\left(\frac{D}{F}\right) - r \quad \textbf{(5.6)}$$

where r is the risk-free rate. For our example, the risk-free rate (implied by the zero-coupon bond price we use) is 10 percent. The yield on the debt is 10.35 percent, so the credit spread is 35 basis points. Not surprisingly, the credit spread falls as the value of the debt rises. The logarithm of D/F in Equation (5.6) is multiplied by $-[1/(T - t)]$. An increase in the value of debt increases D/F, but since the logarithm of D/F is multiplied by a negative number, the credit spread falls.

With debt, the most we can receive at maturity is par. As time to maturity lengthens, it becomes more likely that we will receive less than par. However, if the value of the debt is low enough to start with, there is more of a chance that the value of the debt will be higher as the debt reaches maturity if time to maturity is longer. Consequently, if the debt is highly rated, the spread widens as time to maturity gets longer. For sufficiently risky debt, the spread can narrow as time to maturity gets longer. This is shown in Figure 5-2.

Helwege and Turner (1999) show that credit spreads widen with time to maturity for low-rated public debt, so that even low-rated debt is not risky enough to lead to credit spreads that narrow with time to maturity. It is important to note that credit spreads depend on interest rates. The expected value of the firm at maturity increases with the risk-free rate, so there is less risk that the firm will default. As a result, credit spreads narrow as interest rates increase.

Finding Firm Value and Firm Value Volatility

Suppose a firm, Supplier Inc., has one large debt claim. The firm sells a plant to a very risky third party, In-The-Mail Inc., and instead of receiving cash it receives a promise from In-The-Mail Inc. that it will pay $100 million in five years. We want to value this debt claim. If V_{t+5} is the value of In-The-Mail Inc. at maturity of the debt, we know that the debt pays $F - \text{Max}(F - V_{t+5}, 0)$ or $V_{t+5} - \text{Max}(V_{t+5} - F, 0)$. If it were possible to trade claims on the value of In-The-Mail Inc., pricing the debt would be straightforward. We could simply compute the value of a put on In-The-Mail Inc. with the appropriate exercise price. In general, we cannot directly trade a portfolio of securities that represents a claim to the whole firm; In-The-Mail Inc.'s debt is not a traded security.

The fact that a firm has some nontraded securities creates two problems. First, we cannot observe firm value directly and, second, we cannot trade the firm to hedge a claim whose value depends on the value of the firm. We can solve both problems. Remember that with Merton's model the only random variable that affects the value of claims

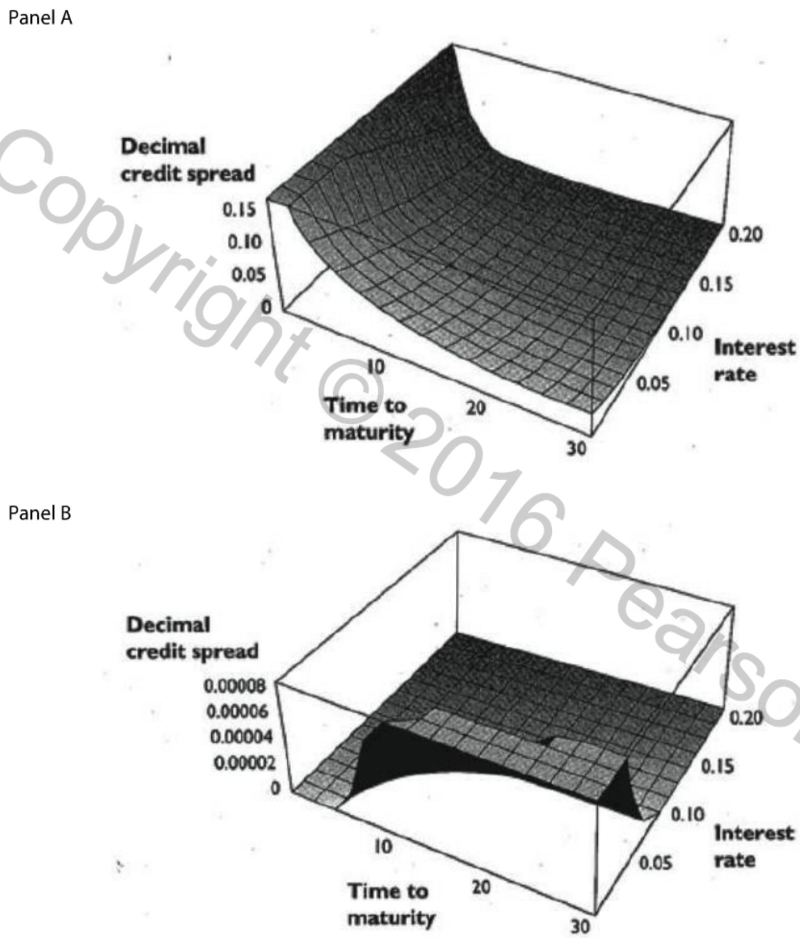


FIGURE 5-2 Credit spreads, time to maturity, and interest rates.

Panel A has firm value of \$50 million, volatility of 20 percent, and debt with face value of \$150 million. Panel B differs only in that firm value is \$200 million.

on the firm is the total value of the firm. Since equity is a call option on firm value, it is a portfolio consisting of δ units of firm value plus a short position in the risk-free asset. The return on equity is perfectly correlated with the return on the value of the firm for small changes in the value of the firm because a small change in firm value of ΔV changes equity by $\delta\Delta V$. We can therefore use equity and the risk-free asset to construct a portfolio that replicates the firm as a whole, and we can deduce firm value from the value of traded claims on the value of the firm. To do that, however, we need to estimate the δ of equity.

To compute the δ of equity from Merton's formula, we need to know firm value V, the volatility of firm value, the promised debt payment, the risk-free interest rate, and the maturity of the debt. If we have this information, computing delta is straightforward. Otherwise, we can estimate these variables using information we have. If we have an estimate of δ and we know the value of the firm's equity, then we can solve for firm value and the volatility of firm value as long as we know the promised debt payment and the maturity of the debt. This is because, in this case, we have two unknowns, firm volatility and firm value, and two equations, the Merton equation for the value of equity and the Merton equation for the equity's delta. We can solve these equations to find firm volatility and the value of the firm. Having these values, we can then solve for the value of the debt. (In practical applications, the Merton model is often used for more complicated capital structures. In this case, the promised debt payment is an estimate of the amount of firm value over some period of time such that if firm value falls below that amount, the firm will be in default and have to file for bankruptcy.)

Suppose we do not know δ. One way to find δ is as follows. We assume In-The-Mail Inc. has traded equity and that its only debt is the debt it owes to Supplier Inc. The value of the debt claim in million dollars is $D(V, 100, t + 5, t)$. The value of a share is \$14.10 and there are 5 million shares. The value of the firm's equity is therefore \$70.5 million. The interest rate is 10 percent per year. Consequently, in million dollars, we have:

$$S(V, 100, t + 5, t) = 70.5 \quad \textbf{(5.7)}$$

We know that equity is a call option on the value of the firm, so that $S(V, 100, t + 5, t) = c(V, 100, t + 5, t)$. We cannot trade V because it is the sum of the debt we own and the value of equity. Since the return to V is perfectly correlated with the return on equity, we can form a dynamic portfolio strategy that pays V_{t+5} at $t + 5$. We will see the details of the strategy later, but for now we assume it can be done.

If we know V, we can use Equation (5.7) to obtain the firm's implied volatility, but the equation has two unknowns: V and the volatility of the value of the firm, σ. To solve for the two unknowns, we have to find an additional equation so that we have two equations and two unknowns. Suppose that there are options traded on the firm's equity. Suppose further that a call option on one share with exercise price of $10 and maturity of one year is worth $6.72. We could use the option pricing formula to get the volatility of equity and deduce the volatility of the firm from the volatility of equity. After all, the option and the equity both depend on the same random variable, the value of the firm.

The difficulty with this is that the Black-Scholes formula does not apply to the call option in our example because it is a call option on equity, which itself is an option when the firm is levered. We therefore have an option on an option, or what is called a compound option. The Black-Scholes formula applies when equity has constant volatility, but the equity in our example cannot have constant volatility if firm value has constant volatility. For a levered firm where firm value has constant volatility, equity is more volatile when firm value is low than when it is high, so that volatility falls as firm value increases. This is because, in percentage terms, an increase in firm value has more of an impact on equity when the value of equity is extremely low than when it is extremely high—even though the equity's δ increases as firm value increases.

A compound call option gives its holder the right to buy an option for a given exercise price. Geske (1979) derives a pricing formula for a compound option that we can use. Geske's formula assumes that firm value follows a lognormal distribution with constant volatility. Therefore, if we know the value of equity, we can use Geske's formula to obtain the value of firm volatility. This formula is presented in Technical Box 5-1, Compound Option Formula.

Pricing the Debt of In-The-Mail Inc.

We now have two equations that we can use to solve for V and σ: the equation for the value of equity (the Black-Scholes formula), and the equation for the value of an option on equity (the compound option formula of Technical Box 5-1). These two equations have only two unknowns. We proceed by iteration.

Suppose we pick a firm value per share of $25 and a volatility of 50 percent. With this, we find that equity should be worth $15.50 and that the call price should be $6.0349. Consequently, the value of equity is too high and the value of the call is too low. Reducing the assumed firm value reduces both the value of equity and the value of the call, so that it brings us closer to the value of equity we want but farther from the value of the call we want. We therefore need to change both firm value and some other variable, taking advantage of the fact that equity and a call option on equity have different greeks. Reducing our assumed firm value reduces the value of equity as well as the value of the call, and increases volatility, which increases the value of equity and the value of the call. In this case, we find an option value of $7.03 and a value of equity of $13.30. Now, the value of equity is too low and the value of the option is too high. This suggests that we have gone too far with our reduction in firm value and increase in volatility. A value of the firm of $21 per share and a volatility of 68.36 percent yield the right values for equity and for the option on equity. Consequently, the value of the debt per share is $6.90. The value of the firm is therefore $105 million. It is divided between debt of $34.5 million and equity of $70.5 million.

Once we have the value of the firm and its volatility, we can use the formula for the value of equity to create a portfolio whose value is equal to the value of the firm and thus can replicate dynamically the value of the firm just using the risk-free asset and equity. Remember that the value of the firm's equity is given by a call option on the value of the firm. Using the Black-Scholes formula, we have:

$$S(V, F, T, t) = VN(d) - P_t(T)FN(d - \sigma\sqrt{T-t})$$
$$d = \frac{\ln(V/P_t(T)F)}{\sigma\sqrt{T-t}} + \frac{1}{2}\sigma\sqrt{T-t} \qquad \textbf{(5.8)}$$

Inverting this formula, the value of the firm is equal to:

$$V = \left(\frac{1}{N(d)}\right)S(V, F, T, t) + P_t(T)F\left(\frac{N(d) - \sigma\sqrt{T-t}}{N(d)}\right) \qquad \textbf{(5.9)}$$

Note that we know all the terms on the right-hand side of this equation. Hence, an investment of $1/N(d)$ of the firm's equity and of $N(d - \sigma\sqrt{T-t})F/N(d)$ units of the zero-coupon bond is equal to the value of the firm per share. Adjusting this portfolio dynamically over time insures that we have V(T) at maturity of the debt. We can scale this portfolio so that it pays off the value of the firm per share. With our example, the portfolio that pays off the value of the firm per share has an investment of 1.15 shares, and an investment in zero-coupon bonds worth $7.91.

BOX 5-1 Compound Option Formula

A compound call option gives its holder the right to buy an option on an option for a given exercise price. Since equity is an option on firm value, an option on the stock of a levered firm is a compound option. Geske (1979) provides a formula for a compound option to value a call on the equity of a levered firm. Geske assumes that firm value follows the same distribution as the stock price in the Black-Scholes formula: firm value has constant volatility and the logarithm of firm value is normally distributed.

Let V be the value of the firm and F be the face value of the debt per share. We define T′ as the maturity date of the option on equity and T the maturity of the debt, where T′ < T. With this, the option holder receives equity at T′ if the option is in the money. Let K be the option exercise price. In exchange for paying K at date T′, the option holder receives equity which is a call on firm value. Using our notation for equity, the value of this call at T′ is S(V, F, T, T′). If S(V, F, T, T′) exceeds K, the option is exercised.

With this notation, the value of the compound option is:

$$Ve^{-d(T-t)}N_2(a_1,b_1;[(T'-t)/(T-t)]^{0.5})$$
$$-Fe^{-r(T-t)}N_2(a_2,b_2;[(T'-t)/(T-t)]^{0.5})-e^{-r(T'-t)}KN(a_2)$$

$$a_1=\frac{\ln(V/V^*)+(r-d+\sigma^2/2)(T'-t)}{\sigma(T'-t)^{0.5}};a_2=a_1-\sigma(T'-t)^{0.5}$$

$$b_1=\frac{\ln(V/F)+(r-d+\sigma^2/2)(T-t)}{\sigma(T-t)^{0.5}};b_2=b_1-\sigma(T-t)^{0.5}$$

V^* is such that, for $t = T'$,

$$V^*e^{-d(T-t)}N(b_1)-Fe^{-r(T-1)}N(b_2)-K=0$$

where $N_2(\alpha, \beta, \rho)$ denotes the cumulative bivariate normal distribution evaluated at α and β for two random variables, each with zero mean and unit standard deviation, that have a correlation coefficient of ρ. The bivariate normal distribution is the distribution followed by two random variables that are jointly normally distributed. The dividend rate is d; it is assumed that dividends are a constant fraction of firm value.

The intuition that we acquired about the determinants of the value of a call option works for compound call options. The value of the compound call option increases when the value of the firm increases, falls when the face value of the debt rises, increases when the time to maturity of the debt rises, increases when the risk-free interest rate rises, increases when the variance of the firm rises, falls as the exercise price rises, and increases as the time to expiration of the call rises.

Subordinated Debt

In principle, subordinated debt receives a payment in the event of bankruptcy only after senior debt has been paid in full. Consequently, when a firm is in poor financial condition, subordinated debt is unlikely to be paid in full and is more like an equity claim than a debt claim. In this case, an increase in firm volatility makes it more likely that subordinated debt will be paid off and hence increases the value of subordinated debt. Senior debt always falls in value when firm volatility increases.

To understand the determinants of the value of the subordinated debt, consider then the case where the senior debt and the subordinated debt mature at the same date. F is the face value of the senior debt and U is the face value of the subordinated debt. Equity is an option on the value of the firm with exercise price U + F since the shareholders receive nothing unless the value of the firm exceeds U + F. Figure 5-3 shows the payoff of

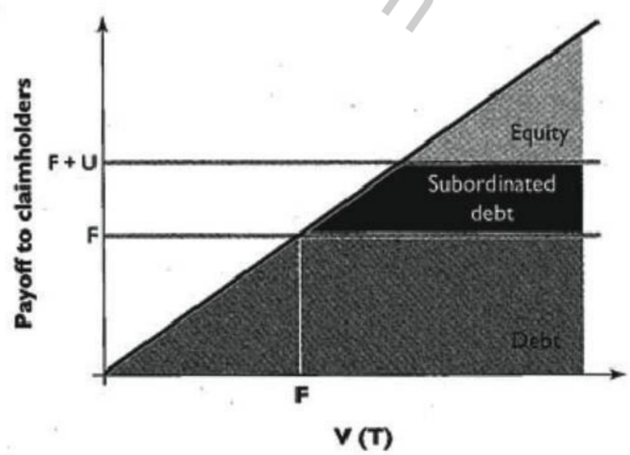


FIGURE 5-3 Subordinated debt payoffs.

Both debt claims are zero-coupon bonds. They mature at the same time. The subordinated debt has face value U and the senior debt has face value F. Firm value is V(T).

subordinated debt as a function of the value of the firm. In this case, the value of the firm is:

$$V = D(V, F, T, t) + SD(V, U, T, t) + S(V, U + F, T, t) \qquad \textbf{(5.10)}$$

D denotes senior debt, SD subordinated debt, and S equity. The value of equity is given by the call option pricing formula:

$$S(V, U + F, T, t) = c(V, U + F, T, t) \qquad \textbf{(5.11)}$$

By definition, shareholders and subordinated debt holders receive collectively the excess of firm value over the face value of the senior debt, F, if that excess is positive. Consequently, they have a call option on V with exercise price equal to F. This implies that the value of the senior debt is the value of the firm V minus the value of the option held by equity and subordinated debt holders:

$$D(V, F, T, t) = V - c(V, F, T, t) \qquad \textbf{(5.12)}$$

Having priced the equity and the senior debt, we can then obtain the subordinated debt by subtracting the value of the equity and of the senior debt from the value of the firm:

$$\begin{aligned} SD(V, U, T, t) &= V - c(V, F + U, T, t) \\ &\quad - [V - c(V, F, T, t)] \\ &= c(V, F, T, t) - c(V, F + U, T, t) \end{aligned} \qquad \textbf{(5.13)}$$

With this formula, the value of subordinated debt is the difference between the value of an option on the value of the firm with exercise price F + U and an option on the value of the firm with exercise price F. Consider a firm with value of $120 million. It has junior debt maturing in five years with face value of $50 million and senior debt maturing in five years with face value of $100 million. The interest rate is 10 percent and the volatility is 20 percent. In this case, we have F = $100 million and U = $50 million. The first call option is worth $60.385 million. It is simply the equity in the absence of subordinated debt, which we computed before. The second call option is worth $36.56 million. The value of the subordinated debt is therefore $60.385M − $36.56M = $23.825 million. Using our formula for the credit spread, we find that the spread on subordinated debt is 4.83 percent, which is more than 10 times the spread on senior debt of 35 basis points.

The fact that the value of subordinated debt corresponds to the difference between the value of two options means that an increase in firm volatility has an ambiguous effect on subordinated debt value. As shown in Figure 5-4, an increase in firm volatility can increase the value of subordinated debt. Subordinated debt is a portfolio that has a long position in a call option that increases in value with volatility and a short position in a call option that becomes more costly as volatility increases. If the subordinated debt is unlikely to pay off, the short position in the call is economically unimportant. Consequently, subordinated debt is almost similar to equity and its value is an increasing function of the volatility of the firm.

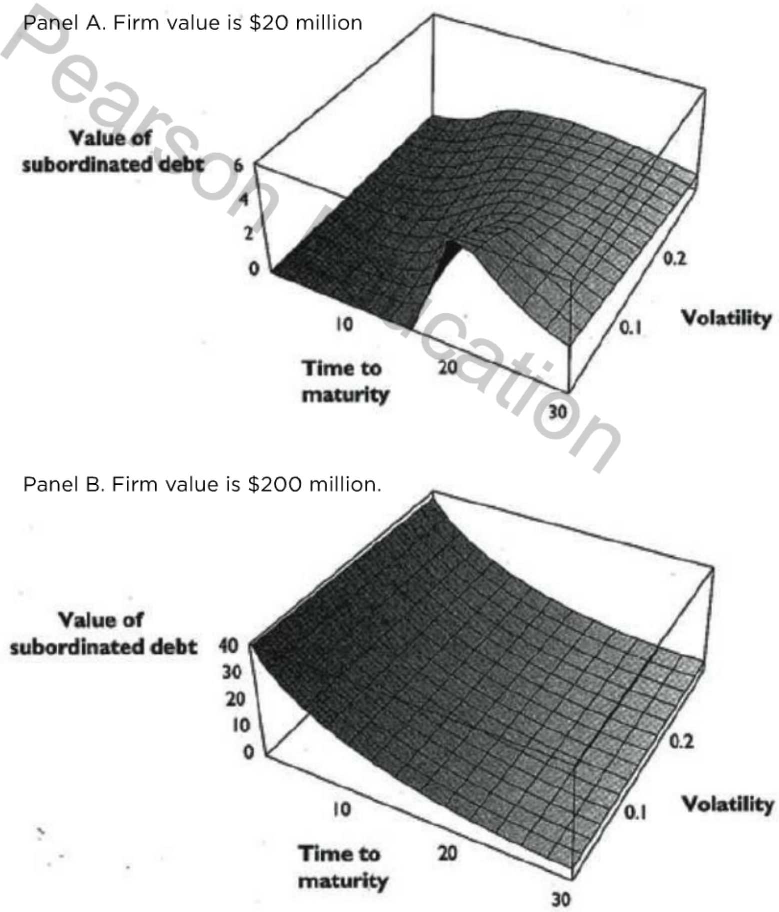


FIGURE 5-4 Subordinated debt, firm value, and volatility.

We consider subordinated debt with face value of $50 million and senior debt with face value of $100 million. The two debt issues mature in five years. The risk-free rate is 10 percent.

Alternatively, if the firm is unlikely to ever be in default, then the subordinated debt is effectively like senior debt and inherits the characteristics of senior debt.

The value of subordinated debt can fall as time to maturity decreases. If firm value is low, the value of the debt increases as time to maturity increases because there is a better chance that it will pay something at maturity. If firm value is high and debt has low risk, it behaves more like senior debt.

Similarly, a rise in interest rates can increase the value of subordinated debt. As the interest rate rises, the value of senior debt falls so that what is left for the subordinated debt holders and equity increases. For low firm values, equity gets little out of the interest rate increase because equity is unlikely to receive anything at maturity, so the value of subordinated debt increases. For high firm values, the probability that the principal will be paid is high, so the subordinated debt is almost risk-free, and its value necessarily falls as the interest rate increases.

The Pricing of Debt When Interest Rates Change Randomly

Unanticipated changes in interest rates can affect debt value through two channels. First, an increase in interest rates reduces the present value of promised coupon payments absent credit risk, and hence reduces the value of the debt. Second, an increase in interest rates can affect firm value. Empirical evidence suggests that stock prices are negatively correlated with interest rates. Hence, an increase in interest rates generally reduces the value of debt both because of the sensitivity of debt to interest rates and because on average it is associated with an adverse shock to firm values. When we want to hedge a debt position, we therefore have to take into account the interaction between interest rate changes and firm value changes.

We consider the pricing of risky debt when the spot interest rate follows the Vasicek model. The change in the spot interest rate over a period of length Δt is:

$$\Delta r_t = \lambda(k - r_t)\Delta t + \sigma_r \varepsilon_t \tag{5.14}$$

where r_t is the current spot interest rate and ε_t is a random shock. When λ is positive, the interest rate reverts to a long-run mean of k. With Equation (5.14) describing how the interest rate evolves, the price of a zero-coupon bond at t that pays \$1 at T, $P_t(T)$, is given by the Vasicek model.

Suppose value and interest rate changes are correlated. Shimko, Tejima, and van Deventer (1993) show that with these interest rate dynamics the value of risky debt is:

$$\begin{aligned}
D(V, r, F, t, T) &= V - VN(h_1) + FP_t(T)N(h_2) \\
Q &= (T-t)\left(\sigma^2 + \frac{\sigma_r^2}{k^2} + \frac{2\rho\sigma\sigma_r}{k}\right) \\
&\quad + (e^{-k(T-t)} - 1)\left(\frac{2\sigma_r^2}{k^3} + \frac{2\rho\sigma\sigma_r}{k^2}\right) \\
&\quad - \frac{\sigma_r^2}{2k^3}(e^{-2k(T-t)} - 1) \\
h_1 &= \frac{\ln\left(\frac{V}{P_t(T)F}\right) + 0.5Q}{\sqrt{Q}} \\
h_2 &= h_1 - \sqrt{Q}
\end{aligned} \tag{5.15}$$

To see how interest rate changes affect the price of debt, we price debt of face value of \$100 million maturing in 5 years on a firm worth \$120 million as we did before. When we assumed a fixed interest rate of 10 percent and a firm volatility of 20 percent, we found that the value of the debt was \$59.615 million at the beginning of the chapter. We choose the parameters of the Vasicek model to be such that, with a spot interest rate of 10 percent, the price of a zero-coupon bond that pays \$1 in 5 years is the same as with a fixed interest rate of 10 percent. This requires us to assume k to be 14.21 percent, the interest rate volatility 10 percent, and the mean reversion parameter 0.25. The volatility of firm value is 20 percent as before, and the correlation between firm value changes and interest rate changes is −0.2. With these assumptions, the debt is then \$57.3011 million.

Figure 5-5 shows how the various parameters of interest rate dynamics affect the price of the debt. We choose a firm value of \$50 million, so that the firm could not repay the debt if it matured immediately. In Panel A, the value of the debt falls as the correlation between firm value and interest rate shocks increases. In this case, firm value is higher when the interest rate is high, so that the impact of an increase in firm value on the value of the debt is more likely to be dampened by a simultaneous interest rate increase. In Panel B, an increase in interest rate volatility and an increase in the speed of mean reversion reduce debt value. With high mean reversion, the interest rate

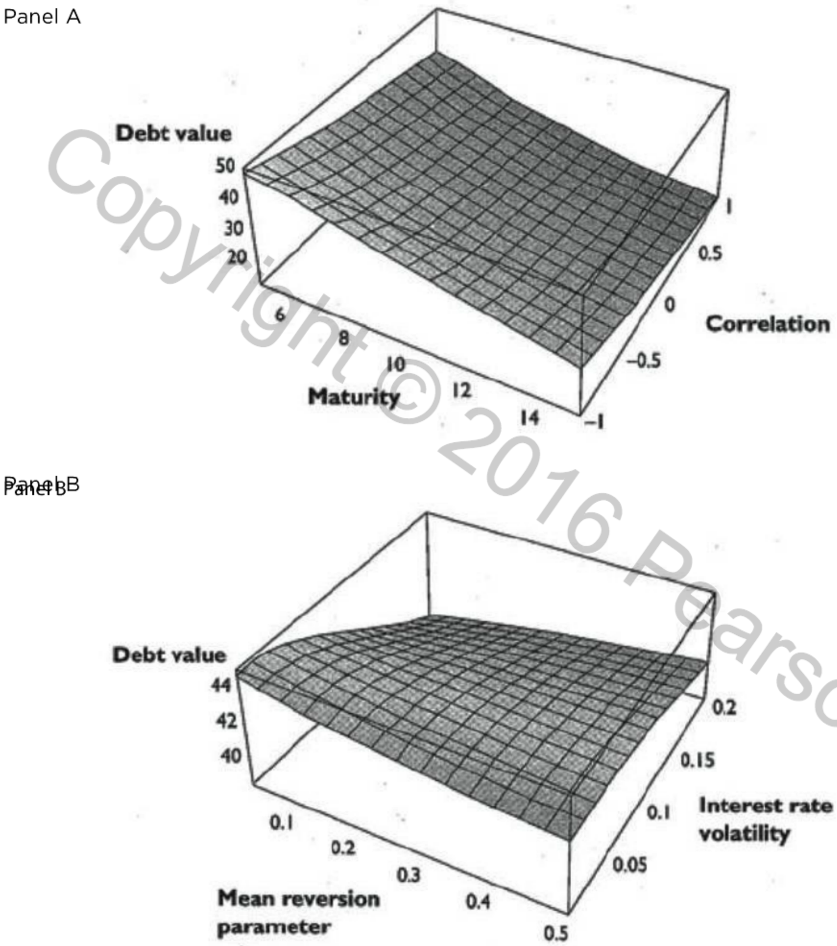


FIGURE 5-5 Subordinated debt, firm value, and volatility.

In the base case, firm value of \$50 million, the promised debt payment is \$100 million, the maturity of the debt is 5 years, the interest rate is set at 10 percent, the spread of mean reversion parameter is 0.25, the volatility of the interest rate is 10 percent, the correlation between firm value and the interest rate is −0.2.

does not diverge for very long from its long-run mean, so that we are closer to the case of fixed interest rates. However, with our assumptions, the long-run mean is higher than the current interest rate.

Figure 5-6 shows that the debt's interest rate sensitivity depends on the volatility of interest rates. At highly volatile interest rates, the value of the debt is less sensitive to changes in interest rate. Consequently, if we were to hedge debt against changes in interest rates, the hedge ratio would depend on the parameters of the dynamics of interest rates.

VaR and Credit Risks

Once we have a pricing model for the valuation of default-risky debt held by the firm, we can incorporate credit risk into the computation of firm-wide risk. Suppose we use the Merton model. Default-risky debt depends on firm value, which itself is perfectly correlated with the debtor's stock price. This makes the debtor's equity one additional risk factor. Suppose we want to compute a VaR measure for the firm, and the firm has just one risky asset: its risky debt. One way is to compute the delta-VaR by transforming the risky debt into a portfolio of the risk-free bond and of the debtor's equity if we are computing a VaR for a short period of time. A second way is to compute the Monte Carlo VaR by simulating equity returns and valuing the debt for these equity returns. If the firm has other assets, we must consider the correlations among the asset returns.

Conceptually, the inclusion of credit risks in computations of VaR does not present serious difficulties. All the difficulties are in the implementation—but they are serious. The complexities of firm capital structures create important obstacles to valuing debt, and often debt is issued by firms with no traded equity. There are thus alternative approaches to debt pricing.

BEYOND THE MERTON MODEL

Corporations generally have many different types of debt with different maturities, and most debt makes coupon payments when not in default. The Merton model approach can be used to price any type of debt. Jones, Mason, and Rosenfeld (1984) test this approach for a panel of firms that include investment-grade firms as well as firms below investment grade. They find that a naive model predicting that debt is riskless works better for investment-grade debt than the Merton model. In contrast, the Merton model works better than the naive model for debt below investment grade. As pointed out by Kim, Ramaswamy, and Sundaresan (1993), however, Merton's

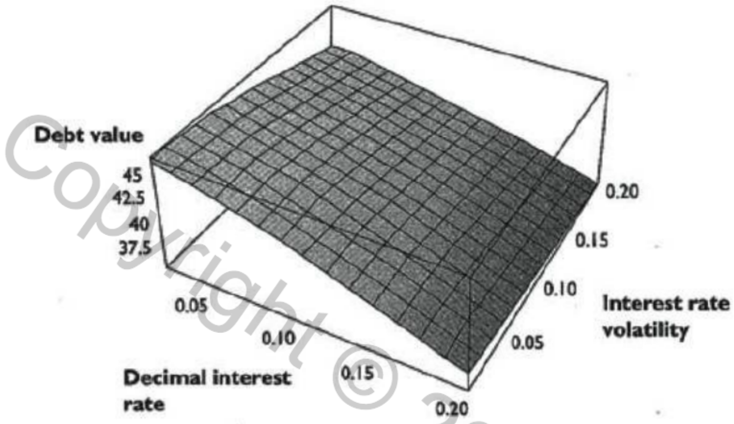


FIGURE 5-6 Interest rate sensitivity of debt.

Firm value is \$50 million, the promised debt payment is \$100 million, the maturity of the debt is 5 years, the speed of mean reversion parameter is 0.25, the correlation between firm value and the interest rate is −0.2.

model fails to predict credit spreads large enough to match empirical data. They point out that from 1926 to 1986, AAA spreads ranged from 15 to 215 basis points, with an average of 77, while BAA spreads ranged from 51 to 787 basis points, with an average of 198. Yet they show that Merton's model cannot generate spreads in excess of 120 basis points.

There are important difficulties in implementing the Merton model when debt makes coupon payments or a firm has multiple debt issues that mature at different dates. Consider the simple case where debt makes one coupon payment u at t' and pays $F + u$ at T. We know how to value the coupon payment u since it is equivalent to a risky zero-coupon debt payment at t'. Valuing the right to $F + u$ to be received at T is harder because it is contingent on the firm paying u at t'. Taking the viewpoint of the equity holders simplifies the analysis. After the equity holders have paid the coupon at t', their claim on the firm is the same as the claim they have in our analysis in the beginning of the chapter, namely a call option on the firm with maturity at T with exercise price equal to the promised payment to the debt holders (which here is $F + u$). Consequently, by paying the coupon at t' the equity holders acquire a call option on the value of the firm at T with exercise price $F + u$. The value of equity at t is the present value of the call option equity holders acquire at t' if they pay the coupon. This is the present value of the payoff of a European call option maturing at t', since they get $\text{Max}(S_{t'} - u, 0)$, where $S_{t'}$ is the value of equity at t'. Hence, at t, the equity holders have a call option on the equity value at t' with exercise price u—they have an option on an option, or a compound option. The value of debt at t is firm value minus a compound option. If we had an additional coupon at t', so that $t' < t'' < T$, we would have to subtract from V at t an option on an option on an option. This creates a considerable computational burden in computing debt value. This burden can be surmounted, but it is not trivial to do so. In practice, this difficulty is compounded by the difficulty that one does not know V because of nontraded debt.

Another difficulty with the Merton model is that default is too predictable. Remember that to obtain prices of debt in that model, we make the Black-Scholes assumptions. We know that with these assumptions firm value cannot jump. As a result, default cannot occur unless firm value is infinitesimally close to the point where default occurs. In the real world, default is often more surprising. For instance, a run on a bank could make its equity worthless even though before the run its equity value was not close to zero.

These problems have led to the development of a different class of models that take as their departure point a probability of default that evolves over time according to a well-defined process. Under this approach, the probability of default can be positive even when firm value significantly exceeds the face value of the debt—this is the case if firm value can jump. The economics of default are modeled as a black box. Default either happens over an interval of time or it does not. Upon default, the debt holder receives a fraction of the promised claim. The recovery rate is the fraction of the principal recovered in the event of default. This recovery rate can be random or certain.

Let's look at the simplest case and assume that the process for the probability of default is not correlated with the interest rate process, and recovery in the event of default is a fixed fraction of the principal amount, θ, which does not depend on time. The bond value next period is $D_{t+\Delta t} + u$ if the bond is not in default, where u is the coupon. If the debt is in default, its value is θF. In the absence of arbitrage opportunities, the bond price today, D_t, is simply the expected value of the bond next period computed using risk-neutral probabilities discounted at the risk-free rate. Using q as the risk-neutral probability of default, it must be the case that:

$$D_t = P_t(t + \Delta t)\,[q\theta F + (1 - q)(D_{t+\Delta t} + u)] \quad \textbf{(5.16)}$$

In this equation, the value of the nondefaulted debt today depends on the value of the nondefaulted debt tomorrow. To solve this problem, we therefore start from the last period. In the last period, the value of the debt is equal to the principal amount plus the last coupon payment, F + u, or to θF. We then work backward to the next-to-the-last period, where we have:

$$D_{T-\Delta t} = P_{T-\Delta t}(T)[q\theta F + (1 - q)(F + u)] \quad \textbf{(5.17)}$$

If we know q and θ, we can price the debt in the next-to-the-last period and continue to keep working backward to get the debt value.

How can we obtain the probability of default and the amount recovered in the event of default? If the firm has publicly traded debt, we can infer these parameters from the price of the debt. Alternatively, we can infer risk-neutral probabilities of default and recovery rates from spreads on debt with various ratings.

Different applications of this approach can allow for random recovery rates as well as for correlations between recovery rates and interest rates or correlations between default rates and interest rates. The empirical evidence shows that this approach works well to price swap spreads and bank subordinated debt.

CREDIT RISK MODELS

There are several differences between measuring the risk of a portfolio of debt claims and measuring the risk of a portfolio of other financial assets. First, because credit instruments typically do not trade on liquid markets where we can observe prices, we cannot generally rely on historical data on individual credit instruments to measure risk. Second, the distribution of returns differs. We cannot assume that continuously compounded returns on debt follow a normal distribution. The return of a debt claim is bounded by the fact that investors cannot receive more than the principal payment and the coupon payments. In statistical terms, this means that the returns to equity are generally symmetric, while the returns to debt are skewed—unless the debt is deeply discounted.

The third difference is that firms often have debt issued by creditors with no traded equity. A fourth difference is that typically debt is not marked to market in contrast to traded securities. When debt is not marked to market, a loss is recognized only if default takes place. Consequently, when debt is not marked to market, a firm must be able to assess the probability of default and the loss made in the event of default.

A number of credit risk models resolve some of the difficulties associated with debt portfolios. Some models focus only on default and on the recovery in the event of default. The most popular model of this type is CreditRisk+, from Credit Suisse Financial Products. It is based on techniques borrowed from the insurance industry for the modeling of extreme events. Other models are based on the marked-to-market value of debt claims. CreditMetrics™ is a risk model built on the same principles as those of RiskMetrics™. The purpose of this risk model is to provide the distribution of the value of a portfolio of debt claims, which leads to a VaR measure for the portfolio. The KMV model is in many ways quite similar to the CreditMetrics™ model, except that it makes direct use of the Merton model in computing the probability of default.[1] All these models can be used to compute the risk of portfolios that include other payoffs besides those of pure debt contracts. For example, they can include swap contracts. As a result, these models talk about estimating the risk of obligors—all those who have legal obligations to the firm—rather than debtors.

To see how we can use the Merton model for default prediction, remember that it assumes that firm value is lognormally distributed with constant volatility and that the firm has one zero-coupon debt issue. If firm value exceeds the face value of debt at maturity, the firm is not in default. We want to compute the probability that firm value will be below the face value of debt at maturity of the debt because we are interested in forecasting the likelihood of a default. To compute this probability, we have to know the expected rate of return of the firm since the higher that expected rate of return, the less likely it is that the firm will be in default. Let m be the expected rate of return of the firm value. In this case, the probability of default is simply:

[1] Oldrich A. Vasicek, Credit Valuation, KMV Corporation, 1984.

$$\text{Probability of default} = N\left(\frac{\ln(F) - \ln(V) - \mu(T-t) + 0.5\sigma^2(T-t)}{\sigma\sqrt{T-t}}\right) \quad \textbf{(5.18)}$$

where N denotes the cumulative normal distribution, F is the face value of the debt, V the value of the firm, T the maturity date of the debt, and σ the volatility of the rate of change of V.

Consider the case where a firm has value of $120 million, debt with face value of $100 million and maturity of five years, the expected rate of change of firm value is 20 percent, the volatility is 20 percent, and the interest rate is 5 percent. The probability that the firm will default is 0.78 percent.

Figure 5-7 shows how the probability of default is related to volatility and firm value. As volatility increases, the probability of default increases. It can be substantial even for large firm values compared to the face value of the debt when volatility is high.

We use the same approach to compute the firm's expected loss if default occurs. The loss if default occurs is often called loss given default or LGD. Since default occurs when firm value is less than the face value of the debt, we have to compute the expected value of V given that it is smaller than F. The solution is:

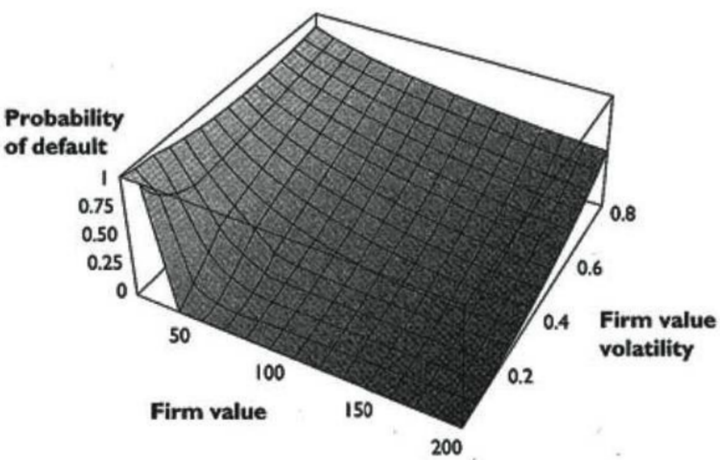


FIGURE 5-7 Probability of default.

The firm value has debt with face value of $100 million due in five years. The firm's expected rate of return is 20 percent.

$$\text{Expected loss} = FN\left(\frac{\ln(F) - \ln(V) - \mu(T-t) + 0.5\sigma^2(T-t)}{\sigma\sqrt{T-t}}\right) - Ve^{\mu(T-t)}N\left(\frac{\ln(F) - \ln(V) - \mu(T-t) - 0.5\sigma^2(T-t)}{\sigma\sqrt{T-t}}\right) \quad \textbf{(5.19)}$$

The expected loss is $100,614.

Figure 5-8 shows how the expected loss depends on firm value and its volatility.

CreditRisk+

CreditRisk+ allows only two outcomes for each firm over the risk measurement period: default and no default. If default occurs, the creditor experiences a loss of fixed size. The probability of default for an obligor depends on its rating, the realization of K risk factors, and the sensitivity of the obligor to the risk factors. The risk factors are common across all obligors, but sensitivity to the risk factors differs across obligors. Defaults across obligors covary only because of the K risk factors. Conditional on the risk factors, defaults are uncorrelated across obligors.

The conditional probability of default for an obligor is the probability of default given the realizations of the risk factors, while the unconditional probability of default is the probability obtained if we do not know the realizations of the risk factors. For example, if there is only one risk factor, say, macroeconomic activity, we would expect the conditional probability of default to be higher when macroeconomic activity is poorer. The unconditional probability of default in this case is the probability when we do not know whether macroeconomic activity is poor or not.

If $p_i(x)$ is the probability of default for the ith obligor conditional on the realizations of the risk factors, and x is the vector of risk factor realizations, the model specifies that:

$$p_i(x) = \pi_{G(i)}\left(\sum_{k=1}^{K} x_k w_{ik}\right) \quad \textbf{(5.20)}$$

where $\pi_{G(i)}$ is the unconditional probability of default for obligor i given that it belongs to grade G. A natural choice of the grade of an obligor would be its public debt rating if it has one. Often, obligors may not have a rating, or the rating of the company may not reflect the

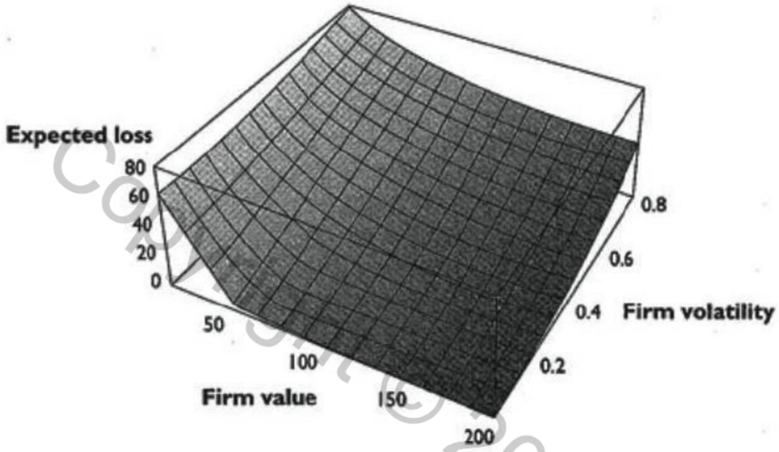


FIGURE 5-8 The expected loss and its dependence on volatility.

The firm has value of $120 million, the face value of the debt is $100 million, the expected rate of change of firm value is 20 percent, and the interest rate is 10 percent.

riskiness of the debt. Bank debt has different covenants than public debt, which makes it easier for banks to intervene when the obligor becomes riskier. As a result, bank debt is less risky than otherwise comparable public debt. The bank internal evaluation system often grades loans on a scale from one to ten. A bank internal grading system could be used to grade obligors.

The risk factors can take only positive values and are scaled so that they have a mean of one. The model also assumes that the risk factors follow a specific statistical distribution (the gamma distribution). If the *k*th risk factor has a realization above one, this increases the probability of default of firm *i* in proportion to the obligor's exposure to that risk factor measured by w_{ik}.

Once we have computed the probability of default for all the obligors, we can get the distribution of the total number of defaults in the portfolio. The relevant distribution is the distribution of losses. The model expresses the loss upon default for each loan in standardized units. A standardized unit could be $1 million. The exposure to the *i*th obligor, $v(i)$, would be an exposure of $v(i)$ standardized units. A mathematical function gives the unconditional probability of a loss of *n* standardized units for each value of *n*. We can also get the volatility of the probability of a loss of *n* standardized units since an unexpectedly high realization of the vector of risk factors will lead to a higher than expected loss.

The CreditRisk+ model is easy to use, largely because of some carefully chosen assumptions about the formulation of the probability of default and the distribution of the risk factors. The statistical assumptions of the model are such that an increase in the volatility of the risk factor has a large impact on the tail of the distribution of the risk factor. Gordy (2000) provides simulation evidence on the CreditRisk+ model. His base case has 5,000 obligors and a volatility of the risk factor of 1. The distribution of grades for obligors is structured to correspond to the typical distribution for a large bank according to Federal Reserve Board statistics. He assumes that the loss upon default is equal to 30 percent of the loan for all loans. Losses are calculated as a percentage of outstanding loans. For a low-quality portfolio (the model rating is BB), the expected loss is 1.872 percent and its volatility is 0.565 percent. The distribution of losses is skewed, so the median of 0.769 percent is much lower than the expected loss. There is a 0.5 percent probability that the loss will exceed 3.320 percent. As the volatility of the risk factor is quadrupled, the mean and standard deviation of losses are essentially unchanged, but there is a 0.5 percent probability that the loss will exceed 4.504 percent.

CreditMetrics™

J.P. Morgan's CreditMetrics™ offers an approach to evaluate the risk of large portfolios of debt claims on firms with realistic capital structures. To see how the CreditMetrics™ approach works, we start from a single debt claim, show how we can measure the risk of the claim with the approach, and then extend the analysis to a portfolio of debt claims.[2]

Consider a debt claim on Almost Iffy Inc. We would like to measure the risk of the value of the debt claim in one year using VaR. To do that, we need to know the fifth quantile

[2] The CreditMetrics™ Technical Manual, available on RiskMetrics website, analyzes the example used here in much greater detail. The data used here is obtained from that manual.

TABLE 5-1 One-Year Transition Matrix

Initial rating	Rating at year-end (%)							
	AAA	AA	A	BBB	BB	B	CCC	Default
AAA	90.81	8.33	0.68	0.06	0.12	0	0	0
AA	0.70	90.65	7.79	0.64	0.06	0.14	0.02	0
A	0.09	2.27	91.05	5.52	0.70	0.26	0.01	0.06
BBB	0.02	0.33	5.95	86.93	5.30	1.17	0.12	0.18
BB	0.03	0.14	0.67	7.73	80.53	8.84	1.00	1.06
B	0	0.11	0.24	0.43	6.48	83.46	4.07	5.20
CCC	0.22	0	0.22	1.30	2.38	11.24	64.86	19.79

TABLE 5-2 One-Year Forward Zero Curves for Various Rating Classes (%)

Rating class	Year 1	Year 2	Year 3	Year 4
AAA	3.60	4.17	4.73	5.12
AA	3.65	4.22	4.78	5.17
A	3.72	4.32	4.93	5.32
BBB	4.10	4.67	5.25	5.63
BB	5.55	6.02	6.78	7.27
B	6.05	7.02	8.03	8.52
CCC	15.05	15.02	14.03	13.52

of the distribution of the value of the debt claim if we use a 5 percent VaR.

Our first step in using the CreditMetrics™ approach is to figure out a rating class for the debt claim. Say that we decide the claim should have a rating BBB. Almost Iffy's debt could remain at that rating, could improve if the firm does better, or could worsen if default becomes more likely. There is a historical probability distribution that a claim with a BBB rating will move to some other rating. Across claims of all ratings, the rating transition matrix presented in Table 5-1 gives us the probability that a credit will migrate from one rating to another over one year. Such matrices are estimated and made available by rating agencies. For a debt claim rated BBB, there is a 1.17 percent probability that the debt claim will have a B rating next year.

To obtain the distribution of the value of the debt claim, we compute the value we expect the claim to have for each rating in one year. Using the term structure of bond yields for each rating category, we can get today's price of a zero-coupon bond for a forward contract to mature in one year. Table 5-2 provides an example of one-year forward zero curves. The rows of the table give us the one-year forward discount rates that apply to zero-coupon bonds maturing in the following four years.

We assume coupons are promised to be paid exactly in one year and at the end of each of the four subsequent years. Say that the coupon is $6. We can use the forward zero curves to compute the value of the bond for each possible rating next year. For example, using Table 5-2, if the bond migrates to a BB rating, the present value of the

TABLE 5-3 The Value of the Debt Claim Across Rating Classes and Associated Probabilities

Year-end rating	Probability (%)	Bond value plus coupon ($)
AAA	0.02	109.37
AA	0.33	109.19
A	5.95	108.66
BBB	86.93	107.55
BB	5.30	102.02
B	1.17	98.10
CCC	0.12	83.64
Default	0.18	51.13

coupon to be paid two years from now as of next year is $6 discounted at the rate of 5.55 percent.

If the bond defaults, we need a recovery rate, which is the amount received in the event of default as a fraction of the principal. Suppose that the bond is a senior unsecured bond. Using historical data, the recovery rate for this type of bond is 51.13 percent.

We can compute the value of the bond for each rating class next year and assign a probability that the bond will end up in each one of these rating classes. Table 5-3 shows the result of such calculations. A typical VaR measure would use the fifth percentile of the bond price distribution, which is a BB rating and a value of $102.02. The mean value of the bond is $107.09, so that the fifth percentile is $5.07 below the mean. Say that the price today is $108. There is a 5 percent chance we will lose at least $5.98.

If we have many claims, we have to make an assumption about the correlations among the various claims. If we know the correlations, we can measure the risk of a portfolio of debt claims using the distribution of the portfolio value. For example, suppose that we have an AAA bond and a B bond whose migration probabilities are independent. That is, knowing that the B bond migrates from a B rating to a BB rating provides no information about the likelihood that the AAA bond will migrate to an AA rating. We compute the probabilities of the transitions for each bond independently and multiply them to obtain the joint probability. Using the transition probability matrix in Table 5-1, we know that the probability of a B bond moving to a BB rating is 6.48 percent and that the probability of an AAA bond moving to an AA rating is 8.33 percent. The probability of these two events happening at the same time is the product of the probabilities of the individual events, $0.0648 \times 0.0833 = 0.0054$, or 0.54 percent.

We can compute the value of the portfolio for that outcome. Once we have computed the value of the portfolio for each possible outcome as well as the probability of each outcome, we have a distribution for the value of the portfolio, and we can compute the fifth percentile to obtain a VaR measure.

If the probabilities of the two bonds moving to a particular rating are not independent, the probability that the B bond moves to BB given that the AAA bond moves to AA is not the product of the probability of the B bond moving to BB and the probability of the AAA bond moving to AA. We have to know the probability that the two bonds will move that way. In other words, we need to know the joint distribution of bond migrations. Note that the values of the portfolio for each possible outcome are the same whether the bond migrations are independent or not. The probabilities of the various outcomes differ depending on the migration correlations. Once we know the probabilities of each outcome, we can compute a distribution for the bond portfolio and again compute its VaR.

The major difficulty using the CreditMetrics™ approach is computing the joint distribution of the migrations of the bonds in the portfolio. One way is to use historical estimates for the joint probabilities of bond migrations. In other words, we could figure how often AAA bonds move to an AA rating and B bonds move to a BB rating. This historical frequency would give us an estimate of the probability that we seek. Once we have the joint probability distribution of transitions for the bonds in the portfolio, we can compute the probability distribution of the portfolio.

In general, though, the historical record of rating migrations will not be enough. The correlation among the rating migrations of two bonds depends on other factors. For example, firms in the same industry are more likely to migrate together. To improve on the historical approach, CreditMetrics™ proposes an approach based on stock returns. Suppose that a firm has a given stock price, and we want to estimate its credit risk. From the rating transition matrices, we know the probability of the firm moving to various ratings. Using the distribution of the stock return, we can compute ranges of returns corresponding to the various ratings—if there is a 5 percent probability of default, a default event corresponds to all stock returns that have a probability of at least 95 percent of being exceeded over the period over which credit risk is computed. Proceeding this way, we can produce stock returns corresponding to the various rating outcomes for each firm whose credit is in the portfolio. The correlations between stock returns can then be used to compute probabilities of various rating outcomes for the credits. For instance, if we have two stocks, we can compute the probability that one stock will be in the BB rating range and the other in the AA rating range.

With a large number of credits, using stock returns to compute the joint distribution of outcomes is time-consuming. To simplify the computation, CreditMetrics™ recommends using a factor model in which stock returns depend on country and industry indices as well as on unsystematic risk. The Credit-Metrics™ technical manual shows how to implement such a model.

The KMV Model

KMV derives default probabilities using the "Expected Default Frequency" for each obligor from an extension of Equation (5.18). KMV computes similar probabilities of default, but assumes a slightly more complicated capital structure in doing so. With KMV's model, the capital structure includes equity, short-term debt, long-term debt, and convertible debt. KMV then solves for the firm value and volatility.

One advantage of the KMV approach is that probabilities of default are obtained using the current equity value, so that any event that affects firm value translates directly into a change in the probability of default. Ratings change only with a lag. Another advantage is that probabilities of default change continually rather than only when ratings change. An increase in equity value reduces the probability of default. In the CreditMetrics™ approach, the value of the firm can change substantially, but the probability of default may remain the same because the firm's rating does not change.

KMV uses an approach inspired by the CAPM to obtain the expected growth of firm values that is required to implement Equation (5.18) and uses a factor model to simplify the correlation structure of firm returns. The assumptions used imply an analytical solution for the loss distribution, so that simulation is not needed to compute a Credit VaR with the KMV model.

Some Difficulties with Credit Portfolio Models

The credit portfolio models just discussed present an important advance in measuring credit risk. At the same time, however, the models as presented have obvious limitations. Some have addressed some of these limitations in implementing the models and other models have been developed trying to avoid some of these limitations, but these models as described are the most popular. Models in their most common implementations do not take into account changes in interest rates or credit spreads. Yet, we know that the value of a portfolio of debt can change both because of changes in default risk and changes in interest rates or credit spreads. Nor do the models do much to take into account current economic conditions. As the economy moves from expansion to recession, the distribution of defaults changes dramatically. For example, default numbers reached a peak in 1991, a recession year, then fell before reaching another peak in 2001, another recession year. Further, the transition correlations increase in recessions. Models that use historical transition matrices cannot take into account changing economic conditions.

CREDIT DERIVATIVES

Credit derivatives are financial instruments whose payoffs are contingent on credit risk realizations. For most credit derivatives, the payoff depends on the occurrence of a "credit event" for a reference entity. Generally, a credit event is (1) failure to make a required payment, (2) restructuring that makes any creditor worse off, (3) invocation of cross-default clause, and (4) bankruptcy. Generally, the required payment or the amount defaulted will have to exceed a minimum value (e.g., $10 million) for the credit event to occur.[3]

[3] For a description of the documentation of a credit derivative trade, see "Inside a credit trade," *Derivatives Strategy,* 1998 (December), 24–28.

Credit derivatives are designed as hedging instruments for credit risks. Consider a bank that has credit exposure to many obligors. Before the advent of loan sales and credit derivatives, banks managed their credit risk mostly through diversification. The problem with that approach to managing credit risk is that it forces a bank to turn down customers with which it has valuable relationships. With a credit derivative, a bank can make a loan to a customer and then hedge part or all of this loan by buying a credit derivative. A highly visible example of such a way to use credit derivatives is discussed in Box 5-2, Citigroup and Enron. Except for a futures contract discussed later, credit derivatives are not traded on exchanges. They are over-the-counter instruments. However, firms can also issue securities publicly that provide them with credit protection.

BOX 5-2 Citigroup and Enron

Citigroup had considerable exposure to Enron in 2000. At that time, Enron was a successful company. It had equity capitalization in excess of $50 billion at the start of the year. Its net income for the year was $979 million. Enron's senior unsecured debt's rating was upgraded, so that it finished the year rated BAA1 by Moody's and BBB+ by Standard and Poor's. Despite all this, Citigroup chose to issue securities for $1.4 billion from August 2000 to May 2001 that effectively hedged Citigroup's exposure to Enron.

Enron's senior unsecured debt kept its ratings until October 2001. In December 2001, Enron's rating was a D; it was bankrupt. Citigroup had a loan exposure of $1.2 billion. It also had some insurance-related obligations. It had collateral for about half of the loan exposure. Most likely, its potential losses were covered by the securities it had issued.

These securities worked as follows. Citibank created a trust. This trust issued five-year notes with fixed interest payments. The proceeds were invested in high-quality debt. If Enron did not go bankrupt, the investors would receive the principal after five years. If Enron did go bankrupt, Citigroup had the right to swap Enron's debt to Citigroup for the securities in the trust.

Citigroup promised a coupon of 7.37 percent. At the time, BAA companies were promising 8.07 percent. However, according to a presentation by Enron's treasurer to Standard and Poor's in 2000, Enron debt was trading above its rating, which led him to pitch a rating of AA. At the same time, he explained that the off-balance sheet debt was not material to Enron.

Source: Daniel Altman, "How Citigroup hedged bets on Enron," *New York Times,* February 8, 2002.

The simplest credit derivative is a put that pays the loss on debt due to default at maturity. A put on the firm value with the same maturity as the debt and with an exercise price equal to the face value of the debt is a credit derivative, called a credit default put, that compensates its holder for the loss due to default if such a loss occurs. The put gives its holder the option to receive the exercise price in exchange of the debt claim. Since the put pays the loss incurred by the debt holder if default takes place, a portfolio of the risky debt and the put option is equivalent to holding default-free debt since the risk of the debt is offset by the purchase of the put. We already priced such a put when we valued In-The-Mail's debt, since that debt was worth risk-free debt minus a put. The holder of In-The-Mail debt was short a put on firm value; by buying the credit default put, the holder of the debt eliminates his credit risk by acquiring an offsetting position in the same put.

The most popular credit derivatives involve swap contracts.[4] One type of contract is called a credit default swap. With this swap, party A makes a fixed annual payment to party B, while party B pays the amount lost if a credit event occurs. For example, if Supplier Inc. in our example wants to get rid of the credit risk of In-The-Mail Inc., it can enter a credit default swap with a bank. It makes fixed payments to the bank. If In-The-Mail Inc. defaults at maturity, the bank pays Supplier Inc. the shortfall due to default (face value minus fair value of the debt at the time of default). The credit default swap for Supplier Inc. is effectively equivalent to buying a credit default put but paying for it by installment. Note that the debt will in general require interest payments before maturity and covenants to be respected. If the obligor fails in fulfilling its obligations under the debt contract, there is a credit event. With the credit event, the default payment becomes due.

The credit default swap can have physical delivery, so that Supplier Inc. would sign over the loan to the bank in the event of a default and would receive a fixed payment. Physical delivery is crucial for loans that do not have a secondary market, but physical delivery of loans involves

[4] Dominic Baldwin, "Business is booming," Credit Risk Special Report, *Risk,* April 1999, 8.

tricky and time-consuming issues. Borrowers often object to having their loans signed over. The settlement period for a credit default swap with physical delivery tends to be longer than for a bond trade. If the transfer of a loan faces objections from the borrower, the settlement might extend beyond 30 days.[5]

A credit default exchange swap requires each party to pay the default shortfall on a different reference asset. Two banks might enter a credit default exchange swap for Bank A to pay the shortfall on debt from Widget Inc. and Bank B to pay the shortfall on debt from In-The-Mail Inc. This way, Bank A reduces its exposure to In-The-Mail Inc. and Bank B reduces its exposure to Widget Inc.

Another popular structure is the total return swap. The party seeking to buy insurance against credit risks receives the return on a risk-free investment and pays the return on an investment with default risk. Suppose a bank, the protection buyer, owns a debt claim worth $80 million today that pays interest of $6 million twice a year in the absence of default for the next five years. In a total return swap, the bank pays what it receives from the debt claim every six months. Assuming that the issuer of the debt claim is not in default, the bank pays $6 million every six months. If the issuer does not pay interest at some due date, then the bank pays nothing. In return, the bank might receive six-month LIBOR on $80 million. At maturity, the obligor repays the principal if he is not in default. Suppose the principal is $100 million. In this case, the bank gets a payment at maturity of $20 million corresponding to the final payment of $100 million minus the initial value of $80 million to the swap counterparty. Or, if the obligor is in default and pays only $50 million, the protection buyer receives from the swap counterparty $80 million minus $50 million, or $30 million. This total return swap guarantees to the bank the cash flows equivalent to the cash flows of a risk-free investment of $80 million.

Pricing a total return swap is straightforward, since it is effectively the exchange of a risky bond for a default-free bond. At initiation, the two bonds have to have the same value.

Another credit derivative is a futures contract. The Chicago Mercantile Exchange Quarterly Bankruptcy Index (QBI) futures contract has been traded since April 1998. The QBI is the total of bankruptcy filings in U.S. courts over a quarter. Most bankruptcies are filed by individuals, which makes the contract appropriate to hedge portfolios of consumer debts, such as credit card debt.

The contract is cash settled and the index level is the number of bankruptcy filings in thousands during the quarter preceding contract expiration. At maturity, the futures price equals the index level. The settlement variation is the change in the futures price times $1,000. The minimum increment in the futures price is $0.025.

CREDIT RISKS OF DERIVATIVES

Since the value of an option is never negative whereas a swap can alternate between positive and negative values, it is not surprising that the credit risks of options are easier to evaluate than the credit risks of swaps.

An option with default risk is called a vulnerable option. At maturity, the holder of an option receives the promised payment only if the writer can make the payment. Suppose the writer is a firm with value V and the option is a European call on a stock with price S. The exercise price of the call is K. Without default risk, the option holder receives Max(S − K, 0) at maturity. With a vulnerable option, the holder receives the promised payment only if it is smaller than V, so that the payoff of the call is:

$$\text{Max}[\text{Min}(V, S - K), 0] \tag{5.21}$$

The current value of the call with default risk is just the present value of this payment. There is no closed-form solution for such an option, but it is not difficult to evaluate its value using a Monte Carlo simulation. The correlation between the value of the firm and the value of the option's underlying asset plays an extremely important role in valuation of the vulnerable option. Suppose that V and S are strongly negatively correlated. In this case, it could be that the option has little value because V is low when the option pays off. If V and S are strongly positively correlated, the option might have almost no credit risk because V is always high when S is high. If an option has credit risk, it becomes straightforward to write an option contract that eliminates that credit risk. The appropriate credit derivative is one that pays the difference between a call without default risk and the vulnerable call:

$$\text{Max}(S - K, 0) - \text{Max}[\text{Min}(V, S - K), 0] \tag{5.22}$$

[5] Dwight Case, "The devil's in the details," *Risk*, August 2000, 26–28.

If we can price the vulnerable call, we can also price the credit derivative that insures the call.

An alternative approach is to compute the probability of default and apply a recovery rate if default occurs. In this case, the option is a weighted average of an option without default risk and of the present value of the payoff if default occurs. Say that the option can default only at maturity and does so with probability p. If default occurs, the holder receives a fraction z of the value of the option. In this case, the value of the option today is $(1 - p)c + pzc$, where c is the value of the option without default risk.

This approach provides a rather simple way to incorporate credit risk in the value of the option when the probability of default is independent of the value of the underlying asset of the option. Say that the probability of default is 0.05 and the recovery rate is 50 percent. In this case, the vulnerable call is worth $(1 - 0.05)c + 0.05 \times 0.5 \times c$, which is 97.5 percent of the value of the call without default risk.

It is often the case that the counterparties to a swap enter margin arrangements to reduce default risk. Nevertheless, swaps can entail default risk for both counterparties. Netting means that the payments between the two counterparties are netted out, so that only a net payment has to be made. We assume that netting takes place and that the swap is treated like a debt claim. If the counterparty due to receive net payments is in default, that counterparty still receives the net payments. This is called the full two-way payment covenant. In a limited two-way payment covenant, the obligations of the counterparties are abrogated if one party is in default.

With these assumptions, the analysis is straightforward when the swap has only one payment. Suppose a market maker enters a swap with a risky credit. The risky credit receives a fixed amount F at maturity of the swap—the fixed leg of the swap—and pays S. S could be the value of equity in an equity swap or could be a floating rate payment determined on some index value at some point after the swap's inception. Let V be the value of the risky credit net of all the debt that is senior to the swap. In this case, the market maker receives S − F in the absence of default risk. This amount can be positive or negative. If the amount is negative, the market maker pays F − S to the risky credit for sure. If the amount is positive, the market maker receives S − F if that amount is less than V.

The swap's payoff to the market maker is:

$$-\text{Max}(F - S, 0) + \text{Max}[\text{Min}(S, V) - F, 0] \quad \textbf{(5.23)}$$

The payment that the risk-free counterparty has to make, F, is chosen so that the swap has no value at inception. Since the risk-free counterparty bears the default risk, in that it may not receive the promised payment, it reduces F to take into account the default risk. To find F, we have to compute the present value of the swap payoff to the market maker.

The first term in Equation (5.23) is minus the value of a put with exercise price F on the underlying asset whose value is S. The second term is the present value of an option on the minimum of two risky assets. Both options can be priced. The correlation between V and S plays a crucial role. As this correlation falls, the value of the put is unaffected, but the value of the option on the minimum of two risky assets falls because for a low correlation it will almost always be the case that one of the assets has a low value.

Swaps generally have multiple payments, however, so this approach will work only for the last period of the swap, which is the payment at T. At the payment date before the last payment date, $T - \Delta t$, we can apply our approach. At $T - 2\Delta t$, however, the market maker receives the promised payment at that date plus the promise of two more payments: the payment at $T - \Delta t$ and the payment at T. The payment at $T - \Delta t$ corresponds to the option portfolio of Equation (5.23), but at $T - \Delta t$ the market maker also has an option on the payment of date T which is itself a portfolio of options. In this case, rather than having a compound option, we have an option on a portfolio of options. Valuation of an option on a portfolio of options is difficult to handle analytically, but as long as we know the dynamics that govern the swap payments in the default-free case as well as when default occurs, we can use Monte Carlo analysis.

SUMMARY

We have developed methods to evaluate credit risks for individual risky claims, for portfolios of risky claims, and for derivatives. The Merton model allows us to price risky debt by viewing it as risk-free debt minus a put written on the firm issuing the debt. The Merton model is practical mostly for simple capital structures with one debt issue that has no coupons. Other approaches to pricing risky debt model the probability of default and then discount the risky cash flows from debt using a risk-neutral distribution of the probability of default. Credit risk models such as the CreditRisk+ model, the CreditMetrics™ model, and the KMV model provide approaches to estimating the VaR for a portfolio of credits. Credit derivatives can be used to hedge credit risks.

Key Concepts

credit default swap
credit event
credit risk
credit spread
CreditMetrics™
KMV model
loss given default (LGD)
obligors
rating transition matrix
recovery rate
vulnerable option

Literature Note

Black and Scholes (1973) had a brief discussion of the pricing of risky debt. Merton (1974) provides a detailed analysis of the pricing of risky debt using the Black-Scholes approach. Black and Cox (1976) derive additional results, including the pricing of subordinated debt and the pricing of debt with some covenants. Geske (1977) demonstrates how to price coupon debt using the compound option approach. Stulz and Johnson (1985) show the pricing of secured debt. Longstaff and Schwartz (1995) extend the model so that default takes place if firm value falls below some threshold. Their model takes into account interest rate risk as well as the possibility that strict priority rules will not be respected. Amin and Jarrow (1992) price risky debt in the presence of interest rates changing randomly using the Black-Scholes approach with a version of the Heath-Jarrow-Morton model.

Duffie and Singleton (1999) provide a detailed overview and extensions of the approaches that model the probability of default. Applications of this approach show that it generally works quite well. Perhaps the easiest application to follow is the work of Das and Tufano (1996). They extract probabilities of default from historical data on changes in credit ratings. Armed with these probabilities and with historical evidence on recovery rates and their correlations with interest rates, they price corporate debt. Instead of using historical estimates of default probabilities and recovery rates, they could have extracted these parameters from credit spreads and their study discusses how this could be done. Jarrow and Turnbull (1995) build an arbitrage model of risky debt where the probability of default can be obtained from the firm's credit spread curve. Jarrow, Lando, and Turnbull (1997) provide a general approach using credit ratings. Another interesting application is Duffie and Singleton (1997), who use this approach to price credit spreads embedded in swaps. Madan and Unal (1998) price securities of savings and loan associations. They show how firm-specific information can be incorporated in the default probabilities.

These various approaches to pricing risky claims have rather weak corporate finance underpinnings. They ignore the fact that firms act differently when their value falls and that they can bargain with creditors. Several recent papers take strategic actions of the debtor into account. Leland (1994) models the firm in an intertemporal setting taking into account taxes and the ability to change volatility. Anderson and Sundaresan (1996) take into account the ability of firms to renegotiate on the value of the debt. Deviations from the doctrine of absolute priority by the courts are described in Eberhart, Moore, and Roenfeldt (1990).

Crouhy, Galai, and Mark (2000) provide an extensive comparative analysis of the CreditRisk+, CreditMetrics™, and KMV models. Gordy (2000) provides evidence on the performance of the first two of these models. Jarrow and Turnbull (2000) critique these models and develop an alternative. Johnson and Stulz (1987) were the first to analyze vulnerable options. A number of papers provide formulas and approaches to analyzing the credit risk of derivatives. Jarrow and Turnbull (1995) provide an approach consistent with the use of the HJM model. Jarrow and Turnbull (1997) show how the approach can be implemented to price the risks of swaps.

The CME-QBT contract is discussed in Arditti and Curran (1998). Longstaff and Schwartz (1995) show how to value credit derivatives.

Spread Risk and Default Intensity Models

6

Learning Objectives

After completing this reading you should be able to:

- Compare the different ways of representing credit spreads.
- Compute one credit spread given others when possible.
- Define and compute the Spread '01.
- Explain how default risk for a single company can be modeled as a Bernoulli trial.
- Explain the relationship between exponential and Poisson distributions.
- Define the hazard rate and use it to define probability functions for default time and conditional default probabilities.
- Calculate the conditional default probability given the hazard rate.
- Calculate risk-neutral default rates from spreads.
- Describe advantages of using the CDS market to estimate hazard rates.
- Explain how a CDS spread can be used to derive a hazard rate curve.
- Explain how the default distribution is affected by the sloping of the spread curve.
- Define spread risk and its measurement using the mark-to-market and spread volatility.

Excerpt is Chapter 7 of Financial Risk Management: Models, History, and Institutions, *by Allan Malz.*

This chapter discusses credit spreads, the difference between risk-free and default-risky interest rates, and estimates of default probabilities based on credit spreads. Credit spreads are the compensation the market offers for bearing default risk. They are not pure expressions of default risk, though. Apart from the probability of default over the life of the security, credit spreads also contain compensation for risk. The spread must induce investors to put up not only with the uncertainty of credit returns, but also liquidity risk, the extremeness of loss in the event of default, for the uncertainty of the timing and extent of recovery payments, and in many cases also for legal risks: Insolvency and default are messy.

Most of this chapter is devoted to understanding the relationship between credit spreads and default probabilities. We provide a detailed example of how to estimate a risk neutral default curve from a set of credit spreads. The final section discusses spread risk and spread volatility.

CREDIT SPREADS

Just as risk-free rates can be represented in a number of ways—spot rates, forward rates, and discount factors—credit spreads can be represented in a number of equivalent ways. Some are used only in analytical contexts, while others serve as units for quoting prices. All of them attempt to decompose bond interest into the part of the interest rate that is compensation for credit and liquidity risk and the part that is compensation for the time value of money:

Yield spread is the difference between the yield to maturity of a credit-risky bond and that of a benchmark government bond with the same or approximately the same maturity. The yield spread is used more often in price quotes than in fixed-income analysis.

i-spread. The benchmark government bond, or a freshly initiated plain vanilla interest-rate swap, almost never has the same maturity as a particular credit-risky bond. Sometimes the maturities can be quite different. The *i*- (or interpolated) spread is the difference between the yield of the credit-risky bond and the linearly interpolated yield between the two benchmark government bonds or swap rates with maturities flanking that of the credit-risky bond. Like yield spread, it is used mainly for quoting purposes.

z-spread. The *z*- (or zero-coupon) spread builds on the zero-coupon Libor curve. It is generally defined as the spread that must be added to the Libor spot curve to arrive at the market price of the bond, but may also be measured relative to a government bond curve; it is good practice to specify the risk-free curve being used. Occasionally the *z*-spread is defined using the forward curve.

If the price of a τ-year credit-risky bond with a coupon of c and a payment frequency of h (measured as a fraction of a year) is $p_{\tau,h}(c)$, the *z*-spread is the constant z that satisfies

$$p_{\tau,h}(c) = ch\sum_{i=1}^{\frac{\tau}{h}} e^{-(r_{ih}+z)ih} + e^{-(r_\tau+z)\tau}$$

ignoring refinements due to day count conventions.

Asset-swap spread is the spread or quoted margin on the floating leg of an asset swap on a bond.

Credit default swap spread is the market premium, expressed in basis points, of a CDS on similar bonds of the same issuer.

Option-adjusted spread (OAS) is a version of the *z*-spread that takes account of options embedded in the bonds. If the bond contains no options, OAS is identical to the *z*-spread.

Discount margin is a spread concept applied to floating rate notes. It is the fixed spread over the current (one- or three-month) Libor rate that prices the bond precisely. The discount margin is thus the floating-rate note analogue of the yield spread for fixed-rate bonds. It is sometimes called the *quoted margin.*

Example 6.1 Credit Spread Concepts

Let's illustrate and compare some of these definitions of credit spread using the example of a U.S. dollar-denominated bullet bond issued by Citigroup in 2003, the 4⅞ percent fixed-rate bond maturing May 7, 2015. As of October 16, 2009, this (approximately) $5\frac{200}{360}$-year bond had a semiannual pay frequency, no embedded options, and at the time of writing was rated Baa 1 by Moody's and A– by S&P. These analytics are provided by Bloomberg's YAS screen.

Its yield was 6.36, and with the nearest-maturity on-the-run Treasury note trading at a yield of 2.35 percent, the yield spread was 401 bps.

The i-spread to the swap curve can be calculated from the five- and six-year swap rates, 2.7385 and 3.0021 percent, respectively. The interpolated 5 200/360-year swap rate is 2.8849 percent, so the i-spread is 347.5 bps.

The z-spread, finally, is computed as the parallel shift to the fitted swap spot curve required to arrive at a discount curve consistent with the observed price, and is equal to 351.8 bps.

To see exactly how the z-spread is computed, let's look at a more stylized example, with a round-number time to maturity and pay frequency, and no accrued interest.

Example 6.2 Computing the z-Spread

We compute the z-spread for a five-year bullet bond with semiannual fixed-rate coupon payments of 7 percent per annum, and trading at a dollar price of 95.00. To compute the z-spread, we need a swap zero-coupon curve, and to keep things simple, we assume the swap curve is flat at 3.5 percent per annum. The spot rate is then equal to a constant 3.470 percent for all maturities.

The yield to maturity of this bond is 8.075 percent, so the i-spread to swaps is $8.075 - 3.50 = 4.575$ percent. The z-spread is the constant z that satisfies

$$0.95 = \frac{0.07}{2} \sum_{i=1}^{5 \cdot 2} e^{-(0.03470+z)\frac{i}{2}} + e^{-(0.03470+z)5}$$

This equation can be solved numerically to obtain $z = 460.5$ bps.

Spread Mark-to-Market

We studied the concept of DV01, the mark-to-market gain on a bond for a one basis point change in interest rates. There is an analogous concept for credit spreads, the "spread01," sometimes called DVCS, which measures the change in the value of a credit-risky bond for a one basis point change in spread.

For a credit-risky bond, we can measure the change in market value corresponding to a one basis point change in the z-spread. We can compute the spread01 the same way as the DV01: Increase and decrease the z-spread by 0.5 basis points, reprice the bond for each of these shocks, and compute the difference.

Example 6.3 Computing the Spread01

Continuing the earlier example, we start by finding the bond values for a 0.5-bps move up and down in the z-spread. The bond prices are expressed per $100 of par value:

$$\frac{0.07}{2} \sum_{i=1}^{5 \cdot 2} e^{-(0.03470+0.04605-0.00005)\frac{i}{2}} + e^{-(0.03470+0.04605-0.00005)5} = 0.950203$$

$$\frac{0.07}{2} \sum_{i=1}^{5 \cdot 2} e^{-(0.03470+0.04605+0.00005)\frac{i}{2}} + e^{-(0.03470+0.04605+0.00005)5} = 0.949797$$

The difference is $95.0203 - 94.9797 = 0.040682$ dollars per basis point per $100 of par value. This would typically be expressed as $406.82 per $1,000,000 of par value. The procedure is illustrated in Figure 6-1.

The spread01 of a fixed-rate bond depends on the initial level of the spread, which in turn is determined by the level and shape of the swap curve, the coupon, and other design features of the bond. The "typical" spread01 for a five-year bond (or CDS) is about $400 per $1,000,000 of bond par value (or notional underlying amount). At very low or high spread levels, however, as seen in Figure 6-2, the spread01 can fall well above or below $400.

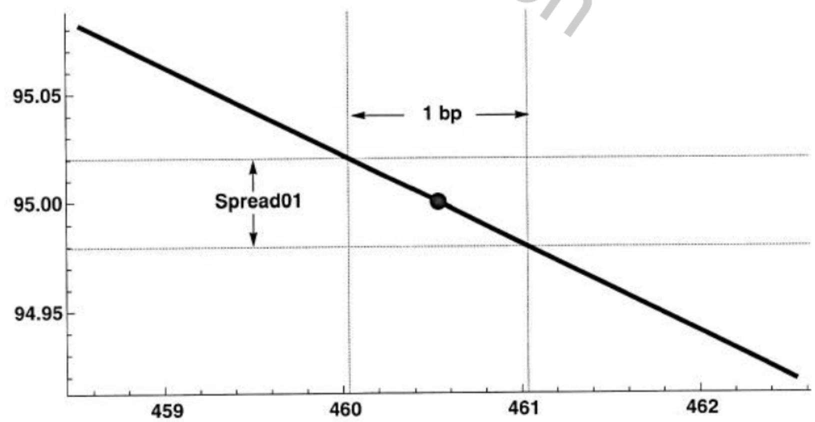


FIGURE 6-1 Computing spread01 for a fixed-rate bond.

The graph shows how spread01 is computed in Example 6.3 by shocking the z-spread up and down by 0.5 bps. The plot displays the value of the bond for a range of z-spreads. The point represents the initial bond price and corresponding z-spread. The vertical grid lines represent the 1 bps spread shock. The horizontal distance between the points on the plot where the vertical grid lines cross is equal to the spread01 per $100 par value.

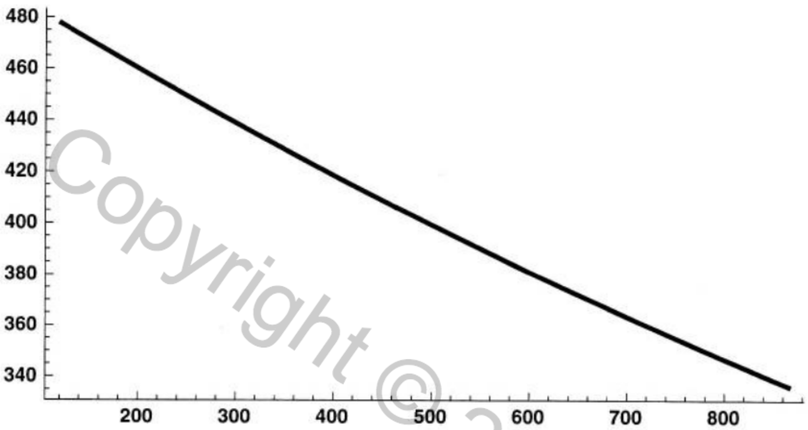


FIGURE 6-2 Spread01 a declining function of spread level.

The graph shows how spread01, measured in dollars per $1,000,000 of bond par value, varies with the spread level. The bond is a five-year bond making semiannual fixed-rate payments at an annual rate of 7 percent. The graph is constructed by permitting the price of the bond to vary between 80 and 110 dollars, and computing the *z*-spread and spread01 at each bond price, holding the swap curve at a constant flat 3.5 percent annual rate.

The intuition is that, as the spread increases and the bond price decreases, the discount factor applied to cash flows that are further in the future declines. The spread-price relationship exhibits convexity; any increase or decrease in spread has a smaller impact on the bond's value when spreads are higher and discount factor is lower. The extent to which the impact of a spread *change* is attenuated by the high *level* of the spread depends primarily on the bond maturity and the level and shape of the swap or risk-free curve.

Just as there is a duration measure for interest rates that gives the proportional impact of a change in rates on bond value, the *spread duration* gives the proportional impact of a spread change on the price of a credit-risky bond. Like duration, spread duration is defined as the ratio of the spread01 to the bond price.

DEFAULT CURVE ANALYTICS

Reduced-form or *intensity models* of credit risk focus on the analytics of default timing. These models are generally focused on practical applications such as pricing derivatives using arbitrage arguments and the prices of other securities, and lead to simulation-friendly pricing and risk measurement techniques. In this section, we lay out the basics of these analytics. Reduced form models typically operate in default mode, disregarding ratings migration and the possibility of restructuring the firm's balance sheet.

Reduced-form models, like the single-factor model of credit risk, rely on estimates of default probability that come from "somewhere else." The default probabilities can be derived from internal or rating agency ratings, or from structural credit models. But reduced form models are most often based on market prices or spreads. These *risk-neutral estimates of default probabilities* can be extracted from the prices of credit-risky bonds or loans, or from credit derivatives such as credit default swaps (CDS). In the next section, we show how to use the default curve analytics to extract default probabilities from credit spread data. In Chapter 7, we use the resulting default probability estimates as an input to models of credit portfolio risk.

Default risk for a single company can be represented as a *Bernoulli trial*. Over some fixed time horizon $\tau = T_2 - T_1$, there are just two outcomes for the firm: Default occurs with probability, π, and the firm remains solvent with probability $1 - \pi$. If we assign the values 1 and 0 to the default and solvency outcomes over the time interval $(T_1, T_2]$, we define a random variable that follows a *Bernoulli distribution*. The time interval $(T_1, T_2]$ is important: The Bernoulli trial doesn't ask "does the firm ever default?," but rather, "does the firm default over the next year?"

The mean and variance of a Bernoulli-distributed variate are easy to compute. The expected value of default on $(T_1, T_2]$ is equal to the default probability π, and the variance of default is $\pi(1 - \pi)$.

The Bernoulli trial can be repeated during successive time intervals $(T_2, T_3], (T_3, T_4], \ldots$ We can set each time interval to have the same length τ, and stipulate that the probability of default occurring during each of these time intervals is a constant value π. If the firm defaults during any of these time intervals, it remains defaulted forever, and the sequence of trials comes to an end. But so long as the firm remains solvent, we can imagine the firm surviving "indefinitely," but not "forever."

This model implies that the Bernoulli trials are *conditionally independent*, that is, that the event of default over each future interval $(T_j, T_{j+1}]$ is independent of the event of default over any earlier ($j > i$) interval $(T_i, T_{i+1}]$. This notion of independence is a potential source of confusion. It means that, from the current perspective, if you are told that the firm will survive up to time T_j, but have no idea when thereafter the firm will default, you "restart the clock" from the perspective of time T_j. You have no more or less information bearing on the survival of the firm over $(T_j, T_j + \tau]$ than you did at an earlier time T_i about survival over $(T_i, T_i + \tau]$. This property is also called *memorylessness*.

In this model, the probability of default over some longer interval can be computed from the binomial distribution. For example, if τ is set equal to one year, the probability of survival over the next decade is equal $(1 - \pi)^{10}$, the probability of getting a sequence of 10 zeros in 10 independent Bernoulli trials.

It is inconvenient, though, to use a discrete distribution such as the binomial to model default over time, since the computation of probabilities can get tedious. An alternative is to model the random time at which a default occurs as the first arrival time—the time at which the modeled event occurs—of a *Poisson process.* In a Poisson process, the number of events in any time interval is *Poisson-distributed.* The time to the next arrival of a Poisson-distributed event is described by the *exponential distribution.* So our approach is equivalent to modeling the time to default as an exponentially distributed random variate. This leads to the a simple algebra describing default-time distributions, illustrated in Figure 6-3.

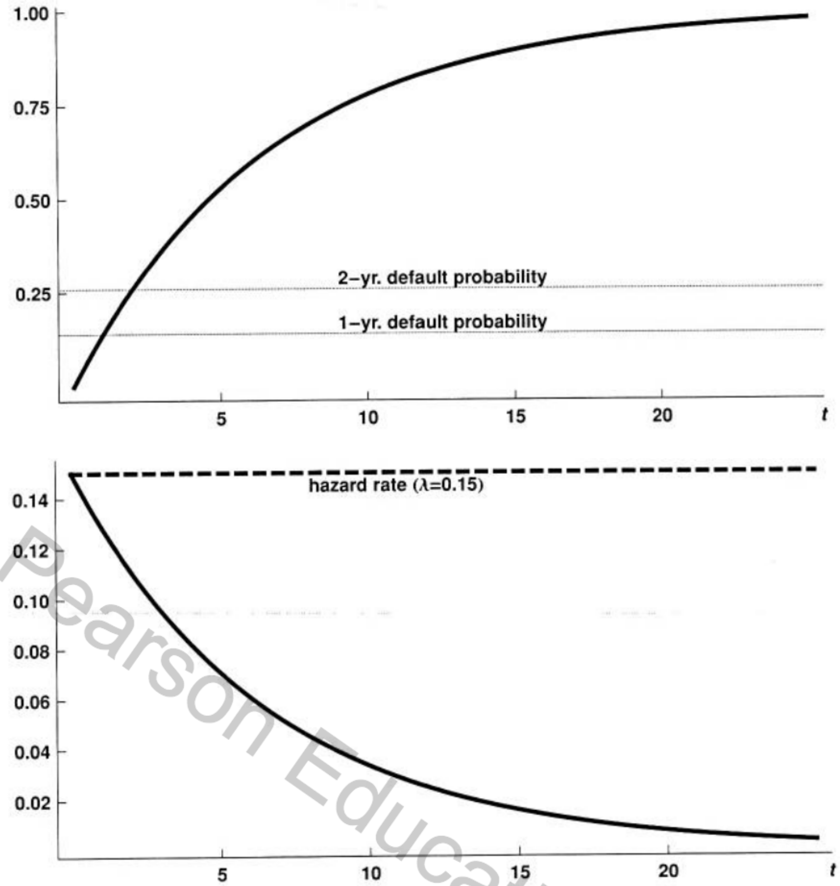


FIGURE 6-3 Intensity model of default timing.

The graphs are plotted from the perspective of time 0 and assume a value $\lambda = 0.15$, as in Example 6.4.
Upper panel: Cumulative default time distribution $1 - e^{-\lambda t}$. The ordinate of each point on the plot represents the probability of a default between time 0 and the time t represented by the abscissa.
Lower panel: Hazard rate λ and marginal default probability $\lambda e^{-\lambda t}$. The ordinate of each point on the plot represents the annual rate at which the probability of a default between time 0 and the time t is changing. The marginal default probability is decreasing, indicating that the one-year probability of default is falling over time.

In describing the algebra of default time distributions, we set $t = 0$ as "now," the point in time from which we are considering different time horizons.

The Hazard Rate

The *hazard rate*, also called the *default intensity*, denoted λ, is the parameter driving default. It has a time dimension, which we will assume is annual.[1] For each future time, the probability of a default over the tiny time interval dt is then

$$\lambda dt$$

and the probability that no default occurs over the time interval dt is

$$1 - \lambda dt$$

[1] In life insurance, the equivalent concept applied to the likelihood of death rather than default is called the *force of mortality.*

In this section, we assume that the hazard rate is a constant, in order to focus on defining default concepts. In the next section, where we explore how to derive risk-neutral default probabilities from market data, we'll relax this assumption and let the hazard rate vary for different time horizons.

Default Time Distribution Function

The *default time distribution function* or *cumulative default time distribution* $F(\tau)$ is the probability of default sometime between now and time t:

$$\mathbf{P}[t^* < t] \equiv F(t) = 1 - e^{-\lambda t}$$

The survival and default probabilities must sum to exactly 1 at every instant t, so the probability of *no* default sometime between now and time t, called the *survival time distribution*, is

$$\mathbf{P}\left[t^* \geq t\right] = 1 - \mathbf{P}\left[t^* < t\right] = 1 - F\left(t\right) = e^{-\lambda t}$$

The survival probability converges to 0 and the default probability converges to 1 as t grows very large: in the intensity model, even a "bullet-proof" AAA-rated company will default eventually. This remains true even when we let the hazard rate vary over time.

Default Time Density Function

The *default time density function* or *marginal default probability* is the derivative of the default time distribution w.r.t. t:

$$\frac{\partial}{\partial t}\mathbf{P}[t^* < t] = F'(t) = \lambda e^{-\lambda t}$$

This is always a positive number, since default risk "accumulates"; that is, the probability of default increases for longer horizons. If λ is small, it will increase at a very slow pace. The survival probability, in contrast, is declining over time:

$$\frac{\partial}{\partial t}\mathbf{P}[t \geq t] = -F'(t) = -\lambda e^{-\lambda t} < 0$$

With a *constant* hazard rate, the marginal default probability is positive but declining, as seen in the lower panel of Figure 6-3. This means that, although the firm is likelier to default the further out in time we look, the rate at which default probability accumulates is declining. This is not necessarily true when the hazard rate can change over time. The default time density is still always positive, but if the hazard rate is rising fast enough with the time horizon, the cumulative default probability may increase at an increasing rather than at a decreasing rate.

Conditional Default Probability

So far, we have computed the probability of default over some time horizon $(0, t)$. If instead we ask, what is the probability of default over some horizon $(t, t + \tau)$ given that there has been no default prior to time t, we are asking about a *conditional default probability.* By the definition of conditional probability, it can be expressed as

$$\mathbf{P}[t^* < t + \tau \mid t^* > t] = \frac{\mathbf{P}[t^* > t \cap t^* < t + \tau]}{\mathbf{P}[t^* > t]}$$

that is, as the ratio of the probability of the joint event of survival up to time t and default over some horizon $(t, t + \tau)$, to the probability of survival up to time t.

That joint event of survival up to time t and default over $(t, t + \tau)$ is simply the event of defaulting during the discrete interval between two future dates t and $t + \tau$. In the constant hazard rate model, the probability of surviving to time t and then defaulting between t and $t + \tau$ is

$$\begin{aligned}\mathbf{P}[t^* > t \cap t^* < t + \tau] &= F(t + \tau) - F(t)\\ &= 1 - e^{-\lambda(t+\tau)} - (1 - e^{-\lambda t})\\ &= e^{-\lambda t}(1 - e^{-\lambda\tau})\\ &= [1 - F(t)]F(\tau)\\ &= \mathbf{P}[t^* > t]\mathbf{P}[t^* < t + \tau | t^* > t]\end{aligned}$$

We also see that

$$F(\tau) = \mathbf{P}[t^* < t + \tau | t^* > t]$$

which is equal to the unconditional τ-year default probability. We can interpret it as the probability of default over τ years, if we started the clock at zero at time t. This useful result is a further consequence of the memorylessness of the default process.

If the hazard rate is constant over a very short interval $(t, t + \tau)$, then the probability the security will default over the interval, given that it has not yet defaulted up until time t, is

$$\lim_{\substack{\tau \to 0 \\ \tau > 0}} \mathbf{P}[t < t^* < t + \tau \mid t^* > t] = \frac{F'(t)\tau}{1 - F(t)} = \lambda\tau$$

The hazard rate can therefore now be interpreted as the instantaneous conditional default probability.

Example 6.4 Hazard Rate and Default Probability

Suppose $\lambda = 0.15$. The unconditional one-year default probability is $1 - e^{-\lambda} = 0.1393$, and the survival probability is $e^{-\lambda}$ 0.8607. This would correspond to a low speculative-grade credit.

The unconditional two-year default probability is $1 - e^{-2\lambda} = 0.2592$. In the upper panel of Figure 6-3, horizontal grid lines mark the one- and two-year default probabilities. The difference between the two- and one-year default probabilities—the probability of the joint event of survival through the first year and default in the second—is 0.11989. The conditional one-year default probability, given survival through the first year, is the difference between the two probabilities (0.11989), divided by the one-year survival probability 0.8607:

$$\frac{0.11989}{0.8607} = 0.1393$$

which is equal, in this constant hazard rate example, to the unconditional one-year default probability.

RISK-NEUTRAL ESTIMATES OF DEFAULT PROBABILITIES

Our goal in this section is to see how default probabilities can be extracted from market prices with the help of the default algebra laid out in the previous section. As noted, these probabilities are risk neutral, that is, they include compensation for both the loss given default and bearing the risk of default and its associated uncertainties. The default intensity model gives us a handy way of representing spreads. We denote the spread over the risk-free rate on a defaultable bond with a maturity of T by z_T. The constant risk-neutral hazard rate at time T is λ^*_T. If we line up the defaultable securities by maturity, we can define a *spread curve*, that is, a function that relates the credit spread to the maturity of the bond.

Basic Analytics of Risk-Neutral Default Rates

There are two main types of securities that lend themselves to estimating default probabilities, bonds and credit default swaps (CDS). We start by describing the estimation process using the simplest possible security, a credit-risky zero-coupon corporate bond.

Let's first summarize the notation of this section:

p_τ	Current price of a default-free τ-year zero-coupon bond
p^{corp}_τ	Current price of a defaultable τ-year zero-coupon bond
r_τ	Continuously compounded discount rate on the default free bond
z_τ	Continuously compounded spread on the defaultable bond
R	Recovery rate
λ^*_τ	τ-year risk neutral hazard rate
$1 - e^{-\lambda^*_\tau}$	Annualized risk neutral default probability

We assume that there are both defaultable and default-free zero-coupon bonds with the same maturity dates. The issuer's credit risk is then expressed by the discount or price concession at which it has to issue bonds, compared to the that on government bonds, rather than the coupon it has to pay to get the bonds sold. We'll assume there is only one issue of defaultable bonds, so that we don't have to pay attention to seniority, that is, the place of the bonds in the capital structure.

We'll denote the price of the defaultable discount bond maturing in τ years by p^{corp}_τ, measured as a decimal. The default-free bond is denoted p_τ. The continuously compounded discount rate on the default-free bond is the spot rate r_τ, defined by

$$p_\tau = e^{-r_\tau \tau}$$

A corporate bond bears default risk, so it must be cheaper than a risk-free bond with the same future cash flows on the same dates, in this case \$1 per bond in τ years:

$$p_\tau \geq p^{corp}_\tau$$

The continuously compounded τ-year spread on a zero-coupon corporate is defined as the difference between the rates on the corporate and default-free bonds and satisfies:

$$p^{corp}_\tau = e^{-(r_\tau + z_\tau)\tau} = p_\tau e^{-z_\tau \tau}$$

Since $p^{corp}_\tau \leq p_\tau$, we have $z_\tau \geq 0$.

The credit spread has the same time dimensions as the spot rate r_τ. It is the constant exponential rate at which, if there is no default, the price difference between a risky and risk-free bond shrinks to zero over the next τ years.

To compute hazard rates, we need to make some assumptions about default and recovery:

- The issuer can default any time over the next τ years.
- In the event of default, the creditors will receive a deterministic and known recovery payment, but only at the maturity date, regardless of when default occurs. Recovery is a known fraction R of the par amount of the bond (recovery of face).

We'll put all of this together to estimate λ^*_τ, the risk-neutral constant hazard rate over the next τ years. The risk-neutral τ-year default probability is thus $1 - e^{-\lambda^*_\tau \tau}$. Later on, we will introduce the possibility of a time-varying hazard rate and learn how to estimate a term structure from bond or CDS data in which the spreads and default probabilities may vary with the time horizon. The time dimensions of λ^*_τ are the same as those of the spot rate and the spread. It is the conditional default probability over $(0, T)$, that is, the constant annualized probability that the firm defaults over a tiny time interval $t + \Delta t$, given that it has not already defaulted by time t, with $0 < t < T$.

The risk-neutral (and physical) hazard rates have an exponential form. The probability of defaulting over the next instant is a constant, and the probability of defaulting over a discrete time interval is an exponential function of the length of the time interval.

For the moment, let's simplify the setup even more, and let the recovery rate $R = 0$. An investor in a defaultable bond receives either \$1 or zero in τ years. The expected value of the two payoffs is

$$e^{-\lambda^*_\tau \tau} \cdot 1 + (1 - e^{-\lambda^*_\tau \tau}) \cdot 0$$

The expected present value of the two payoffs is

$$e^{-r_\tau \tau}[e^{-\lambda^*_\tau \tau} \cdot 1 + (1 - e^{-\lambda^*_\tau \tau}) \cdot 0]$$

Discounting at the risk-free rate is appropriate because we want to estimate λ^*_τ, the risk-neutral hazard rate. To the extent that the credit-risky bond price and z_τ reflect a risk premium as well as an estimate of the true default probability, the risk premium will be embedded in λ^*_τ, so we don't have to discount by a risky rate.

The risk-neutral hazard rate sets the expected present value of the two payoffs equal to the price of the defaultable bond. In other words, if market prices have adjusted to eliminate the potential for arbitrage, we can solve Equation (6.1) for λ^*_τ:

$$e^{-(r_\tau + z_\tau)\tau} = e^{-r_\tau \tau}[e^{-\lambda^*_\tau \tau} \cdot 1 + (1 - e^{-\lambda^*_\tau \tau}) \cdot 0] \qquad \textbf{(6.1)}$$

to get our first simple rule of thumb: If recovery is zero, then

$$\lambda^*_\tau = z_\tau$$

that is, the hazard rate is equal to the spread. Since for small values of x we can use the approximation $e^x \approx 1 + x$, we also can say that the spread $z_t \approx 1 - e^{-\lambda^*_\tau}$, the default probability.

Example 6.5

Suppose a company's securities have a five-year spread of 300 bps over the Libor curve. Then the risk-neutral annual hazard rate over the next five years is 3 percent, and the annualized default probability is approximately 3 percent. The exact annualized default probability is 2.96 percent, and the five-year default probability is 13.9 percent.

Now let the recovery rate R be a positive number on (0, 1). The owner of the bond will receive one of two payments at the maturity date. Either the issuer does not default, and the creditor receives par (\$1), or there is a default, and the creditor receives R. Setting the expected present value of these payments equal to the bond price, we have

$$e^{-(r_\tau + z_\tau)\tau} = e^{-r_\tau \tau}[e^{-\lambda^*_\tau \tau} + (1 - e^{-\lambda^*_\tau \tau})R]$$

or

$$e^{-z_\tau \tau} = e^{-\lambda^*_\tau \tau} + (1 - e^{-\lambda^*_\tau \tau})R = 1 - (1 - e^{-\lambda^*_\tau \tau})(1 - R)$$

giving us our next rule of thumb: The additional credit-risk discount on the defaultable bond, divided by the LGD, is equal to the τ-year default probability:

$$1 - e^{-\lambda^*_\tau \tau} = \frac{1 - e^{-z_\tau \tau}}{1 - R}$$

We can get one more simple rule of thumb by taking logs in Equation (6.1):

$$-(r_\tau + z_\tau)\tau = -r_\tau \tau + \log[e^{-\lambda^*_\tau \tau} + (1 - e^{-\lambda^*_\tau \tau})R]$$

or

$$z_\tau \tau = -\log[e^{-\lambda^*_\tau \tau} + (1 - e^{-\lambda^*_\tau \tau})R]$$

This expression can be solved numerically for λ^*_τ, or we can use the approximations $e^x \approx 1 + x$ and

$\log(1 + x) \approx x$, so $e^{-\lambda^*_\tau \tau} + (1 - e^{-\lambda^*_\tau \tau})R \approx 1 - \lambda^*_\tau \tau + \lambda^*_\tau \tau R = 1 - \lambda^*_\tau \tau(1 - R)$. Therefore,

$$\log[1 - \lambda^*_\tau \tau(1 - R)] \approx -\lambda^*_\tau \tau(1 - R)$$

Putting these results together, we have

$$z_\tau \tau \approx \lambda^*_\tau \tau(1 - R) \Rightarrow \lambda^*_\tau \approx \frac{z_\tau}{1 - R}$$

The spread is approximately equal to the default probability times the LGD. The approximation works well when spreads or risk-neutral default probabilities are not too large.

Example 6.6

Continuing the example of a company with a five-year spread of 300 bps, with a recovery rate $R = 0.40$, we have a hazard rate of

$$\lambda^*_\tau \approx \frac{0.0300}{1 - 0.4} = 0.05$$

or 5 percent.

So far, we have defined spot hazard rates, which are implied by prices of risky and riskless bonds over different time intervals. But just as we can define spot and forward risk-free rates, we can define spot and forward hazard rates. A forward hazard rate from time T_1 to T_2 is the constant hazard rate over that interval. If $T_1 = 0$, it is identical to the spot hazard rate over $(0, T_2)$.

Time Scaling of Default Probabilities

We typically don't start our analysis with an estimate of the hazard rate. Rather, we start with an estimate of the probability of default π over a given time horizon, based on either the probability of default provided by a rating agency or a model, or on a market credit spread.

These estimates of π have a specific time horizon. The default probabilities provided by rating agencies for corporate debt typically have a horizon of one year. Default probabilities based on credit spreads have a time horizon equal to the time to maturity of the security from which they are derived. The time horizon of the estimated default probability may not match the time horizon we are interested in. For example, we may have a default probability based on a one-year transition matrix, but need a five-year default probability in the context of a longer-term risk analysis.

We can always convert a default probability from one time horizon to another by applying the algebra of hazard rates. But we can also use a default probability with one time horizon directly to estimate default probabilities with longer or shorter time horizons. Suppose, for example, we have an estimate of the one-year default probability π_1. From the definition of a constant hazard rate,

$$\pi_1 = 1 - e^{-\lambda}$$

we have

$$\lambda = \log(1 - \pi_1)$$

This gives us an identity

$$\pi_1 = 1 - e^{-\log(1 - \pi_1)}$$

We can then approximate

$$\pi_t = 1 - (1 - \pi_1)^t$$

Credit Default Swaps

So far, we have derived one constant hazard rate using the prices of default-free and defaultable discount bonds. This is a good way to introduce the analytics of risk-neutral hazard rates, but a bit unrealistic, because corporations do not issue many zero-coupon bonds. Most corporate zero-coupon issues are commercial paper, which have a typical maturity under one year, and are issued by only a small number of highly rated "blue chip" companies. Commercial paper even has a distinct rating system.

In practice, hazard rates are usually estimated from the prices of CDS. These have a few advantages:

Standardization. In contrast to most developed-country central governments, private companies do not issue bonds with the same cash flow structure and the same seniority in the firm's capital structure at fixed calendar intervals. For many companies, however, CDS trading occurs regularly in standardized maturities of 1, 3, 5, 7, and 10 years, with the five-year point generally the most liquid.

Coverage. The universe of firms on which CDS are issued is large. Markit Partners, the largest collector and purveyor of CDS data, provides curves on about 2,000 corporate issuers globally, of which about 800 are domiciled in the United States.

Liquidity. When CDS on a company's bonds exist, they generally trade more heavily and with a tighter bid-offer spread than bond issues. The liquidity of CDS with different maturities usually differs less than that of bonds of a given issuer.

Figure 6-4 displays a few examples of CDS credit curves.

Hazard rates are typically obtained from CDS curves via a bootstrapping procedure. We'll see how it works using a detailed example. We first need more detail on how CDS contracts work. We also need to extend our discussion of the default probability function to include the possibility of time-varying hazard rates. CDS contracts with different terms to maturity can have quite different prices or spreads.

To start, recall that in our simplified example above, the hazard rate was found by solving for the default probability that set the expected present value of the credit spread payments equal to the expected present value of the default loss. Similarly, to find the default probability function using CDS, we set the expected present value of the spread payments by the protection buyer equal to the expected present value of the protection seller's payments in the event of default.

The CDS contract is written on a specific *reference entity*, typically a firm or a government. The contract defines an event of default for the reference entity. In the event of default, the contract obliges the protection seller to pay the protection buyer the par amount of a deliverable bond of the reference entity; the protection buyer delivers the bond. The CDS contract specifies which of the reference entity's bonds are "deliverable," that is, are covered by the CDS.

In our discussion, we will focus on single-name corporate CDS, which create exposure to bankruptcy events of a single issuer of bonds such as a company or a sovereign entity. Most, but not all, of what we will say about how CDS work also applies to other types, such as CDS on credit indexes.

CDS are traded in spread terms. That is, when two traders make a deal, the price is expressed in terms of the spread premium the counterparty buying protection is to pay to the counterparty selling protection. CDS may trade "on spread" or "on basis." When the spread premium would otherwise be high, CDS trade points upfront, that is, the protection buyer pays the seller a market-determined percentage of the notional at the time of the trade, and the spread premium is set to 100 or 500 bps, called "100 running." Prior to the so-called "Big Bang" reform of CDS trading conventions that took place on March 13, 2009, only CDS on issuers with wide spreads traded "points up." The running spread was, in those cases, typically 500 bps. The reformed convention has all CDS trading points up, but with some paying 100 and others 500 bps running.

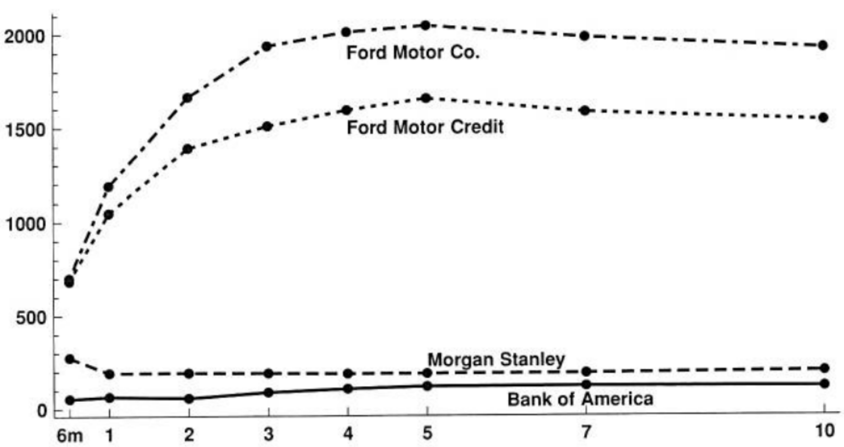


FIGURE 6-4 CDS curves.

CDS on senior unsecured debt as a function of tenor, expressed as an annualized CDS premium in basis points, July 1, 2008.
Source: Bloomberg Financial L.P.

A CDS is a swap, and as such

- Generally, no principal or other cash flows change hands at the initiation of the contract. However, when CDS trade points upfront, a percent of the principal is paid by the protection buyer. This has an impact primarily on the counterparty credit risk of the contract rather than on its pricing, since there is always a spread premium, with no points up front paid, that is equivalent economically to any given market-adjusted number of points upfront plus a running spread. There are generally exchanges of collateral when a CDS contract is created.
- Under the terms of a CDS, there are agreed future cash flows. The protection buyer undertakes to make spread

payments, called the *fee leg*, each quarter until the maturity date of the contract, unless and until there is a default event pertaining to the underlying name on which the CDS is written. The protection seller makes a payment, called the *contingent leg*, only if there is a default. It is equal to the estimated loss given default, that is, the notional less the recovery on the underlying bond.[2]

- The pricing of the CDS, that is, the market-adjusted spread premium, is set so that the expected net present value of the CDS contract is zero. In other words, on the initiation date, the expected present value of the fee leg is equal to that of the contingent leg. If market prices change, the net present value becomes positive for one counterparty and negative for the other; that is, there is a mark-to-market gain and loss.

The CDS contract specifies whether the contract protects the senior or the subordinated debt of the underlying name. For companies that have issued both senior and subordinated debt, there may be CDS contracts of both kinds.

Often, risk-neutral hazard rates are calculated using the conventional assumption about recovery rates that $R = 0.40$. An estimate based on fundamental credit analysis of the specific firm can also be used. In some cases, a risk-neutral estimate is available based on the price of a *recovery swap* on the credit. A recovery swap is a contract in which, in the event of a default, one counterparty will pay the actual recovery as determined by the settlement procedure on the corresponding CDS, while the other counterparty will pay a fixed amount determined at initiation of the contract. Subject to counterparty risk, the counterparty promising that fixed amount is thus able to substitute a fixed recovery rate for an uncertain one. When those recovery swap prices can be observed, the fixed rate can be used as a risk-neutral recovery rate in building default probability distributions.

[2] There is a procedure for cash settlement of the protection seller's contingent obligations, standard since April 2009 as part of the "Big Bang," in which the recovery amount is determined by an auction mechanism. The seller may instead pay the buyer the notional underlying amount, while the buyer delivers a bond, from a list of acceptable or "deliverable" bonds issued by the underlying name rather than make a cash payment. In that case, it is up to the seller to gain the recovery value either through the bankruptcy process or in the marketplace.

Building Default Probability Curves

Next, let's extend our earlier analysis of hazard rates and default probability distributions to accommodate hazard rates that vary over time. We will add a time argument to our notation to indicate the time horizon to which it pertains. The conditional default probability at time t, the probability that the company will default over the next instant, given that it has survived up until time t, is denoted $\lambda(t)$, $t \in [0, \infty)$.

The default time distribution function is now expressed in terms of an integral in hazard rates. The probability of default over the interval $[0, t)$ is

$$\pi_t = 1 - e^{-\int_0^t \lambda(s)ds} \qquad \textbf{(6.2)}$$

If the hazard rate is constant, $\lambda(t) = \lambda$, $t \in [0, \infty)$, then Equation (6.2) reduces to our earlier expression $\pi_t = 1 - e^{-\lambda t}$. In practice, we will be estimating and using hazard rates that are not constant, but also don't vary each instant. Rather, since we generally have the standard CDS maturities of 1, 3, 5, 7, and 10 years available, we will extract 5 piecewise constant hazard rates from the data:

$$\lambda(t) = \begin{Bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ \lambda_5 \end{Bmatrix} \text{ for } \begin{Bmatrix} 0 < t \le 1 \\ 1 < t \le 3 \\ 3 < t \le 5 \\ 5 < t \le 7 \\ 7 < t \end{Bmatrix}$$

The integral from which default probabilities are calculated via Equation (6.2) is then

$$\int_0^t \lambda(s)ds = \begin{Bmatrix} \lambda_1 t \\ \lambda_1 + (t-1)\lambda_2 \\ \lambda_1 + 2\lambda_2 + (t-3)\lambda_3 \\ \lambda_1 + 2\lambda_2 + 2\lambda_3 + (t-5)\lambda_4 \\ \lambda_1 + 2\lambda_2 + 2\lambda_3 + 3\lambda_4 + (t-7)\lambda_5 \end{Bmatrix} \text{ for } \begin{Bmatrix} 0 < t \le 1 \\ 1 < t \le 3 \\ 3 < t \le 5 \\ 5 < t \le 7 \\ 7 < t \end{Bmatrix}$$

Now, let's look at the expected present value of each CDS leg. Denote by s_τ the spread premium on a τ-year CDS on a particular company. The protection buyer will pay the spread in quarterly installments if and only if the credit is still alive on the payment date. The probability of survival up to date t is $1 - \pi_t$, so we can express this expected present value, in dollars per dollar of underlying notional, as

$$\frac{1}{4 \times 10^4} s_\tau \sum_{u=1}^{4\tau} p_{0.25u}(1 - \pi_{0.25u})$$

where p_t is the price of a risk-free zero-coupon bond maturing at time t. We will use a discount curve based on interest-rate swaps. The summation index u takes on integer values, but since we are adding up the present values of quarterly cash flows, we divide u by 4 to get back to time measured in years.

There is one more wrinkle in the fee leg. In the event of default, the protection buyer must pay the portion of the spread premium that accrued between the time of the last quarterly payment and the default date. This payment isn't included in the summation above. The amount and timing is uncertain, but the convention is to approximate it as half the quarterly premium, payable on the first payment date following default. The implicit assumption is that the default, if it occurs at all, occurs midway through the quarter. The probability of having to make this payment on date t is equal to $\pi_t - \pi_{t-0.25}$, the probability of default during the interval $(t - \frac{1}{4}, t]$. This probability is equal to the probability of surviving to time $(t - \frac{1}{4}]$ minus the smaller probability of surviving to time t.

Taking this so-called fee accrual term into account, the expected present value of the fee leg becomes

$$\frac{1}{4 \times 10^4} s_\tau \sum_{u=1}^{4\tau} p_{0.25u}\left[(1 - \pi_{0.25u}) + \frac{1}{2}(\pi_{0.25u} - \pi_{0.25(u-1)})\right]$$

Next, we calculate the expected present value of the contingent leg. If a default occurs during the quarter ending at time t, the present value of the contingent payment is $(1 - R)p_t$ per dollar of notional. We assume that the contingent payment is made on the quarterly cash flow date following the default. The expected present value of this payment is obtained by multiplying this present value by the probability of default during the quarter:

$$(1 - R)p_t(\pi_t - \pi_{t-0.25})$$

The expected present value of the contingent leg is therefore equal to the sum of these expected present values over the life of the CDS contract:

$$(1 - R)\sum_{u=1}^{4\tau} p_{0.25u}(\pi_{0.25u} - \pi_{0.25(u-1)})$$

The fair market CDS spread is the number s_τ that equalizes these two payment streams, that is, solves

$$\frac{1}{4 \times 10^4} s_\tau \sum_{u=1}^{4\tau} p_{0.25u}\left[(1 - \pi_{0.25u}) + \frac{1}{2}(\pi_{0.25u} - \pi_{0.25(u-1)})\right] \quad \textbf{(6.3)}$$
$$= (1 - R)\sum_{u=1}^{4\tau} p_{0.25u}(\pi_{0.25u} - \pi_{0.25(u-1)})$$

Now we're ready to estimate the default probability distribution. To solve Equation (6.3), the market must "have in its mind" an estimate of the default curve, that is the π_t. Of course, it doesn't: The s_τ are found by supply and demand. But once we observe the spreads set by the market, we can infer the π_t by backing them out of Equation (6.3) via a bootstrapping procedure, which we now describe.

The data we require are swap curve interest data, so that we can estimate a swap discount curve, and a set of CDS spreads s_τ on the same name and with the same seniority, but with different terms to maturity. We learned how to generate a swap curve from observation on money-market and swap rates, so we will assume that we can substitute specific numbers for all the discount factors p_t.

Let's start by finding the default curve for a company for which we have only a single CDS spread, for a term, say, of five years. This will result in a single hazard rate estimate. We need default probabilities for the quarterly dates $t = 0.25, 0.50, \ldots, 5$. They are a function of the as-yet unknown hazard rate λ: $\pi_t = 1 - e^{-\lambda * t}$, $t > 0$. Substituting this, the five-year CDS spread, the recovery rate and the discount factors into the CDS valuation function (6.3) gives us

$$\frac{s_\tau}{4 \times 10^4} \sum_{u=1}^{4\tau} p_{0.25u}\left[e^{-\lambda\frac{u}{4}} + \frac{1}{2}\left(e^{-\lambda\frac{u-1}{4}} - e^{-\lambda\frac{u}{4}}\right)\right]$$
$$= (1 - R)\sum_{u=1}^{4\tau} p_{0.25u}\left(e^{-\lambda\frac{u-1}{4}} - e^{-\lambda\frac{u}{4}}\right)$$

with $\tau = 5$. This is an equation in one unknown variable that can be solved numerically for λ.

Example 6.7

We compute a constant hazard rate for Merrill Lynch as of October 1, 2008, using the closing five-year CDS spread of 445 bps. We assume a recovery rate $R = 0.40$. To simplify matters, we also assume a flat swap curve, with a continuously compounded spot rate of 4.5 percent for all maturities, so the discount factor for a cash flow t years in the future is $e^{-0.045t}$. As long as this constant swap rate is reasonably close to the actual swap rate prevailing on October 1, 2008, this has only a small effect on the numerical results.

With $\tau = 5$, $s_\tau = 445$, $R = 0.40$, we have

$$\frac{445}{4 \times 10^4} \sum_{u=1}^{4\cdot 5} e^{0.045\frac{u}{4}}\left[e^{-\lambda\frac{u}{4}} + \frac{1}{2}\left(e^{-\lambda\frac{u-1}{4}} - e^{-\lambda\frac{u}{4}}\right)\right]$$
$$= 0.60\sum_{u=1}^{4\cdot 5} e^{0.045\frac{u}{4}}\left(e^{-\lambda\frac{u-1}{4}} - e^{-\lambda\frac{u}{4}}\right)$$

This equation can be solved numerically to obtain $\lambda = 0.0741688$.

The bootstrapping procedure is a bit more complicated, since it involves a sequence of steps. But each step is similar to the calculation we just carried out for a single CDS spread and a single hazard rate. The best way to explain it is with an example.

Example 6.8

We will compute the default probability curve for Merrill Lynch as of October 1, 2008. The closing CDS spreads on that date for each CDS maturity were

i	τ_i(yrs)	s_{τ_i}(bps/yr)	λ_i
1	1	576	0.09600
2	3	490	0.07303
3	5	445	0.05915
4	7	395	0.03571
5	10	355	0.03416

The table above also displays the estimated forward hazard rates, the extraction of which we now describe in detail. We continue to assume a recovery rate $R = 0.40$ and a flat swap curve, with the discount function $p_t = e^{-0.045t}$.

At each step i, we need quarterly default probabilities over the interval $(0, \tau_i]$ $i = 1, \ldots, 5$, some or all of which will still be unknown when we carry out that step. We progressively "fill in" the integral in Equation (6.2) as the bootstrapping process moves out the curve. In the first step, we find

$$\pi_t = 1 - e^{-\lambda_1 t} \qquad t \in (0, \tau_1]$$

We start by solving for the first hazard rate λ_1. We need the discount factors for the quarterly dates $t = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$, and the CDS spread with the shortest maturity, τ_1. We solve this equation in one unknown for λ_1:

$$\frac{1}{4 \times 10^4} s_{\tau_1} \sum_{u=1}^{4\tau_1} p_{0.25u} \left[e^{-\lambda_1 \frac{u}{4}} + \frac{1}{2}\left(e^{-\lambda_1 \frac{u-1}{4}} - e^{-\lambda_1 \frac{u}{4}} \right) \right]$$
$$= (1-R) \sum_{u=1}^{4\tau_1} p_{0.25u} \left(e^{-\lambda_1 \frac{u-1}{4}} - e^{-\lambda_1 \frac{u}{4}} \right)$$

With $\tau_1 = 1$, $s_{\tau_1} = 576$, and $R = 0.40$, this becomes

$$\frac{576}{4 \times 10^4} \sum_{u=1}^{4} e^{-0.045 \frac{u}{4}} \left[e^{-\lambda_1 \frac{u}{4}} + \frac{1}{2}\left(e^{-\lambda_1 \frac{u-1}{4}} - e^{-\lambda_1 \frac{u}{4}} \right) \right]$$
$$= 0.60 \sum_{u=1}^{4} e^{-0.045 \frac{u}{4}} \left(e^{-\lambda_1 \frac{u-1}{4}} - e^{-\lambda_1 \frac{u}{4}} \right)$$

which we can solve numerically for λ_1, obtaining $\lambda_1 = 0.0960046$. Once the default probabilities are substituted back in, the fee and the contingent legs of the swap are found to each have a fair value of \$0.0534231 per dollar of notional principal protection.

In the next step, we extract λ_2 from the data, again by setting up an equation that we can solve numerically for λ_2. We now need quarterly default probabilities and discount factors over the interval $(0, \tau_2] = (0, 3]$. For any t in this interval,

$$\pi_t = 1 - e^{-\int_0^t \lambda(s)ds} = \begin{Bmatrix} 1 - e^{-\lambda_1 t} \\ 1 - e^{-[\lambda_1 + (t-1)\lambda_2]} \end{Bmatrix} \text{ for } \begin{Bmatrix} 0 < t \le 1 \\ 1 < t \le 3 \end{Bmatrix}$$

The default probabilities for $t \le \tau_1 = 1$ are known, since they use only λ_1.

Substitute these probabilities, as well as the discount factors, recovery rate, and the three-year CDS spread into the expression for CDS fair value to get:

$$\frac{1}{4 \times 10^4} s_{\tau_2} \sum_{u=1}^{4\tau_1} p_{0.25u} \left[e^{-\lambda_1 \frac{u}{4}} + \frac{1}{2}\left(e^{-\lambda_1 \frac{u-1}{4}} - e^{-\lambda_1 \frac{u}{4}} \right) \right]$$
$$+ \frac{1}{4 \times 10^4} s_{\tau_2} e^{-\lambda_1 \tau_1} \sum_{u=4\tau_1+1}^{4\tau_2} p_{0.25u} \left[e^{-\lambda_2 \left(\frac{u}{4} - \tau_1\right)} + \frac{1}{2}\left(e^{\lambda_2 \left[\frac{u-1}{4} - \tau_1\right]} - e^{\lambda_2 \left(\frac{u}{4} - \tau_1\right)} \right) \right]$$
$$= (1-R) \sum_{u=1}^{4\tau_1} p_{0.25u} \left(e^{-\lambda_1 \frac{u-1}{4}} - e^{-\lambda_1 \frac{u}{4}} \right)$$
$$+ (1-R) e^{-\lambda_1 \tau_1} \sum_{u=4\tau_1+1}^{4\tau_2} p_{0.25u} \left(e^{-\lambda_2 \left[\frac{u-1}{4} - \tau_1\right]} - e^{-\lambda_2 \left(\frac{u}{4} - \tau_1\right)} \right)$$

and solve numerically for λ_2.

Notice that the first term on each side of the above equation is a known number at this point in the bootstrapping process, since the default probabilities for horizons of one year or less are known. Once we substitute the known quantities into the above equation, we have

$$\frac{490}{4 \times 10^4} \sum_{u=1}^{4} e^{-0.045 \frac{u}{4}} \left[e^{-0.0960046 \frac{u}{4}} + \frac{1}{2}\left(e^{-0.0960046 \frac{u-1}{4}} - e^{-0.0960046 \frac{u}{4}} \right) \right]$$
$$+ \frac{490}{4 \times 10^4} e^{-0.0960046} \sum_{u=5}^{4 \cdot 3} e^{-0.045 \frac{u}{4}} \left\{ e^{-\lambda_2 \left(\frac{u}{4} - 1\right)} + \frac{1}{2}\left[e^{\lambda_2 \left(\frac{u-1}{4} - 1\right)} - e^{\lambda_2 \left(\frac{u}{4} - 1\right)} \right] \right\}$$
$$= 0.04545$$

$$
\begin{aligned}
&+\frac{490}{4\times10^4}e^{-0.0960046}\sum_{u=5}^{4\cdot3}e^{-0.045\frac{u}{4}}\left\{e^{-\lambda_2\left(\frac{u}{4}-1\right)}+\frac{1}{2}\left[e^{\lambda_2\left(\frac{u-1}{4}-1\right)}-e^{\lambda_2\left(\frac{u}{4}-1\right)}\right]\right\}\\
&=0.60\sum_{u=1}^{4}e^{-0.045\frac{u}{4}}\left(e^{-0.0960045\frac{u-1}{4}}e^{-0.0960046\frac{u}{4}}\right)\\
&+0.60\sum_{u=5}^{4\cdot3}e^{-0.045\frac{u}{4}}\left[e^{\lambda_2\left(\frac{u-1}{4}-1\right)}-e^{\lambda_2\left(\frac{u}{4}-1\right)}\right]\\
&=0.05342+0.60\sum_{u=5}^{4\cdot3}e^{-0.045\frac{u}{4}}\left[e^{\lambda_2\left(\frac{u-1}{4}-1\right)}-e^{\lambda_2\left(\frac{u}{4}-1\right)}\right]
\end{aligned}
$$

which can be solved numerically to obtain $\lambda_2 = 0.0730279$.

Let's spell out one more step explicitly and extract λ_3 from the data. The quarterly default probabilities and discount factors we now need cover the interval $(0, \tau_3) = (0, 5)$. For any t in this interval,

$$
\pi_t = 1-e^{-\int_0^t \lambda(s)ds} = \begin{Bmatrix} 1-e^{-\lambda_1 t} \\ 1-e^{-[\lambda_1+(t-1)\lambda_2]} \\ 1-e^{-[\lambda_1+2\lambda_2+(t-3)\lambda_3]} \end{Bmatrix} \text{ for } \begin{Bmatrix} 0<t\le 1 \\ 1<t\le 3 \\ 3<t\le 5 \end{Bmatrix}
$$

The default probabilities for $t \le \tau_2 = 3$ are known, since they are functions of λ_1 and λ_2 alone, which are known after the second step.

Now we use the five-year CDS spread in the expression for CDS fair value to set up:

$$
\begin{aligned}
&\frac{s_{\tau_3}}{4\times10^4}\sum_{u=1}^{4\tau_1}p_{0.25u}\left[e^{-\lambda_1\frac{u}{4}}+\frac{1}{2}\left(e^{-\lambda_1\frac{u-1}{4}}-e^{-\lambda_1\frac{u}{4}}\right)\right]\\
&+\frac{s_{\tau_3}}{4\times10^4}e^{-\lambda_1\tau_1}\sum_{u=4\tau_1+1}^{4\tau_2}p_{0.25u}\left[e^{-\lambda_2\frac{u}{4}}+\frac{1}{2}\left(e^{-\lambda_2\frac{u-1}{4}}-e^{-\lambda_2\frac{u}{4}}\right)\right]\\
&+\frac{s_{\tau_3}}{4\times10^4}e^{-[\lambda_1\tau_1+\lambda_2(\tau_2-\tau_1)]}\sum_{u=4\tau_2+1}^{4\tau_3}p_{0.25u}\left[e^{-\lambda_3\frac{u}{4}}+\frac{1}{2}\left(e^{-\lambda_3\frac{u-1}{4}}-e^{-\lambda_3\frac{u}{4}}\right)\right]\\
&=(1-R)\sum_{u=1}^{4\tau_1}p_{0.25u}\left(e^{-\lambda_1\frac{u-1}{4}}-e^{-\lambda_1\frac{u}{4}}\right)\\
&+(1-R)e^{-\lambda_1\tau_1}\sum_{u=4\tau_1+1}^{4\tau_2}p_{0.25u}\left(e^{-\lambda_2\frac{u-1}{4}}-e^{-\lambda_2\frac{u}{4}}\right)\\
&+(1-R)e^{-[\lambda_1\tau_1+\lambda_2(\tau_2-\tau_1)]}\sum_{u=4\tau_2+1}^{4\tau_3}p_{0.25u}\left(e^{\lambda_3\frac{u-1}{4}}-e^{\lambda_3\frac{u}{4}}\right)
\end{aligned}
$$

Once again, at this point in the bootstrapping process, since the default probabilities for horizons of three years or less are known, the first two terms on each side of the equals sign are known quantities. And once we have substituted them, we have

$$
\begin{aligned}
&0.10974\\
&+\frac{445}{4\times10^4}e^{-[\lambda_1+2\lambda_2]}\sum_{u=4\tau_2+1}^{4\tau_3}p_{0.25u}\left\{e^{-\lambda_3\left(\frac{u}{4}-1\right)}+\frac{1}{2}\left[e^{\lambda_3\left(\frac{u-1}{4}-1\right)}-e^{\lambda_3\left(\frac{u}{4}-1\right)}\right]\right\}\\
&=0.12083+0.60\sum_{u=4\tau_2+1}^{4\tau_3}p_{0.25u}\left[e^{\lambda_3\left(\frac{u-1}{4}-1\right)}-e^{\lambda_3\left(\frac{u}{4}-1\right)}\right]
\end{aligned}
$$

which can be solved numerically to obtain $\lambda_3 = 0.05915$.

The induction process should now be clear. It is illustrated in Figure 6-5. With our run of five CDS maturities, we repeat the process twice more. The intermediate results are tabulated by step in the table below. Each row in the table displays the present expected value of either leg of the CDS after finding the contemporaneous hazard rate, and the values of the fee and contingent legs up until that step. Note that last period's value of either leg becomes the next period's value of contingent leg payments in previous periods:

i	Either Leg $\to \tau_i$	Fee Leg $\to \tau_{i-1}$	Contingent Leg $\to \tau_{i-1}$
1	0.05342	0.00000	0.00000
2	0.12083	0.04545	0.05342
3	0.16453	0.10974	0.12083
4	0.18645	0.14605	0.16453
5	0.21224	0.16757	0.18645

The CDS in our example did not trade points up, in contrast to the standard convention since 2009. However, Equation (6.3) also provides an easy conversion between pure spread quotes and points up quotes on CDS.

To keep things simple, suppose that both the swap and the hazard rate curves are flat. The swap rate is a continuously compounded r for any term to maturity, and the hazard rate is λ for any horizon. Suppose further that the running spread is 500 bps. The fair market CDS spread will then be a constant s for any term τ. From Equation (6.3), the expected present value of all the payments by the protection buyer must equal the expected present value of loss given default, so the points upfront and the constant hazard rate must satisfy

$$
\begin{aligned}
\frac{\text{points upfront}}{100} &= (1-R)\sum_{u=1}^{4\tau}e^{-r\frac{u}{4}}\left(e^{-\lambda\frac{u-1}{4}}-e^{-\lambda\frac{u}{4}}\right)\\
&-\frac{500}{4\times10^4}\sum_{u=1}^{4\tau}e^{-r\frac{u}{4}}\left[e^{-\lambda\frac{u}{4}}+\frac{1}{2}\left(e^{-\lambda\frac{u-1}{4}}-e^{-\lambda\frac{u}{4}}\right)\right]
\end{aligned}
\qquad \textbf{(6.4)}
$$

Once we have a hazard rate, we can solve Equation (6.4) for the spread from the points upfront, and vice versa.

The Slope of Default Probability Curves

Spread curves, and thus hazard curves, may be upward- or downward-sloping. An upward-sloping spread curve leads to a default distribution that has a relatively flat slope for shorter horizons, but a steeper slope for more

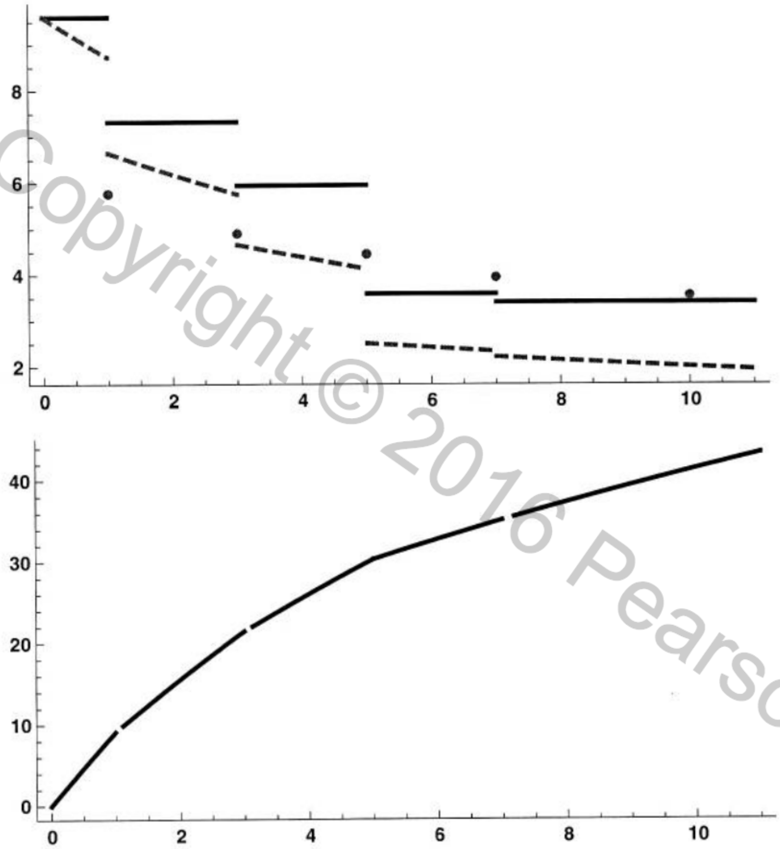


FIGURE 6-5 Estimation of default curves.

Upper panel shows the CDS spreads from which the hazard rates are computed as dots, the estimated hazard rates as a step function (solid plot). The default density is shown as a dashed plot.
Lower panel shows the default distribution. Notice the discontinuities of slope as we move from one hazard rate to the next.

distant ones. The intuition is that the credit has a better risk-neutral chance of surviving the next few years, since its hazard rate and thus unconditional default probability has a relatively low starting point. But even so, its marginal default probability, that is, the conditional probability of defaulting in future years, will fall less quickly or even rise for some horizons.

A downward-sloping curve, in contrast, has a relatively steep slope at short horizons, but flattens out more quickly at longer horizons. The intuition here is that, if the firm survives the early, "dangerous" years, it has a good chance of surviving for a long time.

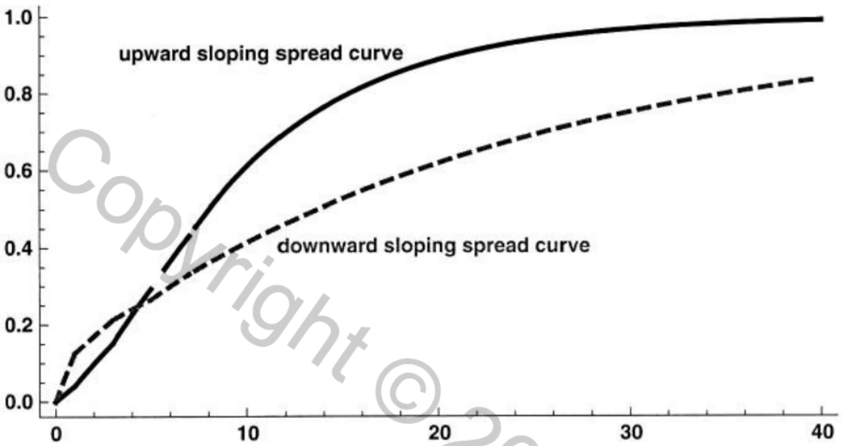
An example is shown in Figure 6-6. Both the upward- and downward-sloping spread curves have a five-year spread of 400 basis points. The downward-sloping curve corresponds to an unconditional default probability that is higher than that of the upward-sloping curve for short horizons, but significantly lower than that of the upward-sloping curve for longer horizons.

Spread curves are typically gently upward sloping. If the market believes that a firm has a stable, low default probability that is unlikely to change for the foreseeable future, the firm's spread curve would be flat if it reflected default expectations only. However, spreads also reflect some compensation for risk. For longer horizons, there is a greater likelihood of an unforeseen and unforeseeable change in the firm's situation and a rise in its default probability. The increased spread for longer horizons is in part a risk premium that compensates for this possibility.

Downward-sloping spread curves are unusual, a sign that the market views a credit as distressed, but became prevalent during the subprime crisis. Figure 6-7 displays an example typical for financial intermediaries, that of Morgan Stanley (ticker MS), one of the five large broker-dealers not associated with a large commercial bank within a bank holding company during the period preceding the crisis. (The other large broker-dealers were Bear Stearns, Lehman Brothers, Merrill Lynch, and Goldman Sachs.) Before the crisis, the MS spread curve was upward-sloping. The level of spreads was, in retrospect, remarkably low; the five-year CDS spread on Sep. 25, 2006 was a mere 21 basis points, suggesting the market considered a Morgan Stanley bankruptcy a highly unlikely event.

The bankruptcy of Lehman Brothers cast doubt on the ability of any of the remaining broker-dealers to survive, and also showed that it was entirely possible that senior creditors of these institutions would suffer severe credit losses. Morgan Stanley in particular among the remaining broker-dealers looked very vulnerable. Bear Stearns had already disappeared; Merrill Lynch appeared likely to be acquired by a large commercial bank, Bank of America; and Goldman Sachs had received some fresh capital and was considered less exposed to credit losses than its peers.

By September 25, 2008, the five-year CDS spread on MS senior unsecured debt had risen to 769 basis points. Its 6-month CDS spread was more than 500 basis points



Graph displays cumulative default distributions computed from these CDS curves, in basis points:

Term	Upward	Downward
1	250	800
3	325	500
5	400	400
7	450	375
10	500	350

A constant swap rate of 4.5 percent and a recovery rate of 40 percent were used in extracting hazard rates.

FIGURE 6-6 Spread curve slope and default distribution.

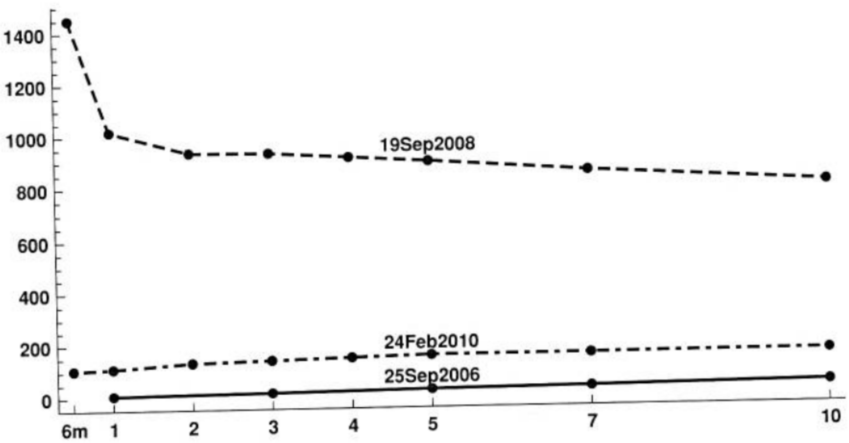


FIGURE 6-7 Morgan Stanley CDS curves, select dates Morgan Stanley senior unsecured CDS spreads, basis points.

Source: Bloomberg Financial L.P.

higher at 1,325 bps. At a recovery rate of 40 percent, this corresponded to about a 12 percent probability of bankruptcy over the next half-year. The one-year spread was over 150 times larger than two years earlier. Short selling of MS common equity was also widely reported, even after the company announced on September 25 that the Federal Reserve Board had approved its application to become a bank holding company.

One year later, the level of spreads had declined significantly, though they remained much higher than before the crisis. On Feb. 24, 2010, the MS five-year senior unsecured CDS spread was 147 basis points, and the curve was gently upward-sloping again.

SPREAD RISK

Spread risk is the risk of loss from changes in the pricing of credit-risky securities. Although it only affects credit portfolios, it is closer in nature to market than to credit risk, since it is generated by changes in prices rather than changes in the credit state of the securities.

Mark-to-Market of a CDS

We can use the analytics of the previous section to compute the effect on the mark-to-market value of a CDS of a change in the market-clearing premium. At initiation, the mark-to-market value of the CDS is zero; neither counterparty owes the other anything. If the spread increases, the premium paid by the fixed-leg counterparty increases. This causes a gain to existing fixed-leg payers, who in retrospect got into their positions cheap, and a loss to the contingent-leg parties, who are receiving less premium than if they had entered the position after the spread widening. This mark-to-market effect is the spread01 of the CDS.

To compute the mark-to-market, we carry out the same steps needed to compute the spread01 of a fixed-rate bond. In this case, however, rather than increasing and decreasing one spread number, the z-spread, by 0.5 bps, we carry out a parallel shift up and down of the entire CDS curve by 0.5 bps. This is similar to the procedure we carried out in computing DV01 for a default-free bond, in which we shifted the entire spot curve up or down by 0.5 bps.

For each shift of the CDS curve away from its initial level, we recompute the hazard rate curve, and with the shocked hazard rate curve we then recompute the value of the CDS. The difference between the two shocked values is the spread01 of the CDS.

Spread Volatility

Fluctuations in the prices of credit-risky bonds due to the market assessment of the value of default and credit transition risk, as opposed to changes in risk-free rates, are expressed in changes in credit spreads. Spread risk therefore encompasses both the market's expectations of credit risk events and the credit spread it requires in equilibrium to put up with credit risk. The most common way of measuring spread risk is via the *spread volatility* or "spread vol," the degree to which spreads fluctuate over time. Spread vol is the standard deviation—historical or expected—of changes in spread, generally measured in basis points per day.

Figure 6-8 illustrates the calculations with the spread volatility of five-year CDS on Citigroup senior U.S. dollar-denominated bonds. The enormous range of variation and potential for extreme spread volatility is clear from the top panel, which plots the spread levels in basis points. The center panel shows daily spread changes (also in bps). The largest changes occur in the late summer and early autumn of 2008, as the collapses of Fannie Mae and Freddie Mac, and then of Lehman, shook confidence in the solvency of large intermediaries, and Citigroup in particular. Many of the spread changes during this period are extreme outliers from the average—as measured by the root mean square—over the entire period from 2006 to 2010.

The bottom panel plots a rolling daily spread volatility estimate, using the EWMA weighting scheme. The calculations are carried out using the recursive form Equation, with the root mean square of the first 200 observations of spread changes as the starting point. The volatility is expressed in basis points per day. A spread volatility of,

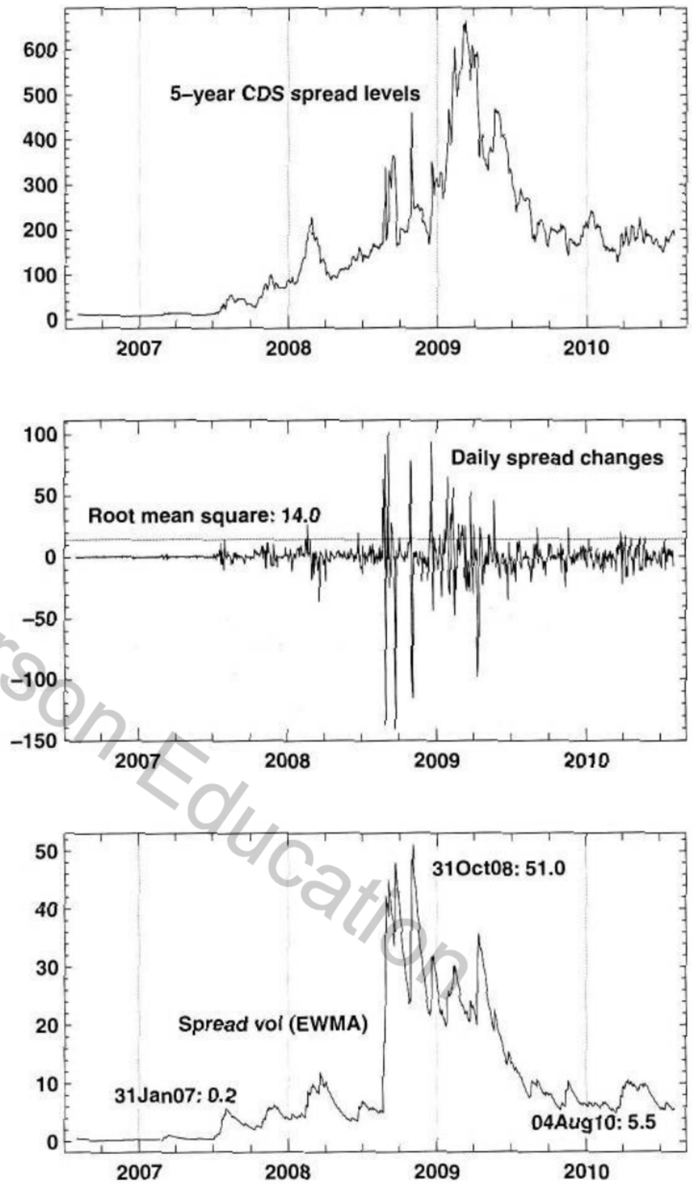


FIGURE 6-8 Measuring spread volatility: Citigroup spreads 2006–2010.

Citigroup 5-year CDS spreads, August 2, 2006, to September 2, 2010. All data expressed in bps.
Source: Bloomberg Financial L.P.
Upper panel: Spread levels.
Center panel: Daily spread changes.
Lower panel: Daily EWMA estimate of spread volatility at a daily rate.

say, 10 bps, means that, if you believe spread changes are normally distributed, you would assign a probability of about 2 out of 3 to the event that tomorrow's spread level is within ±10 bps of today's level. For the early part of the period, the spread volatility is close to zero, a mere quarter of a basis point, but spiked to over 50 bps in the fall of 2008.

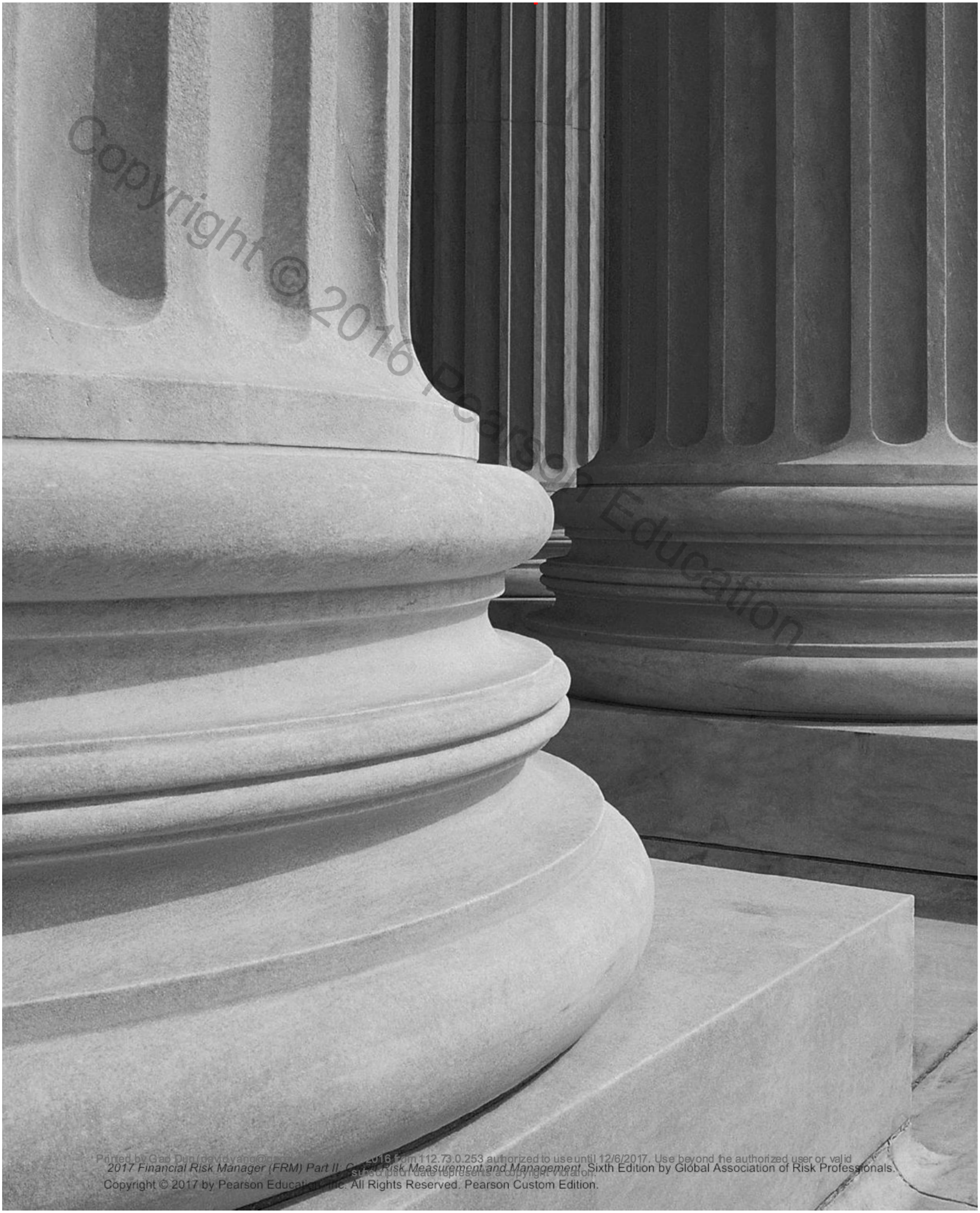
Further Reading

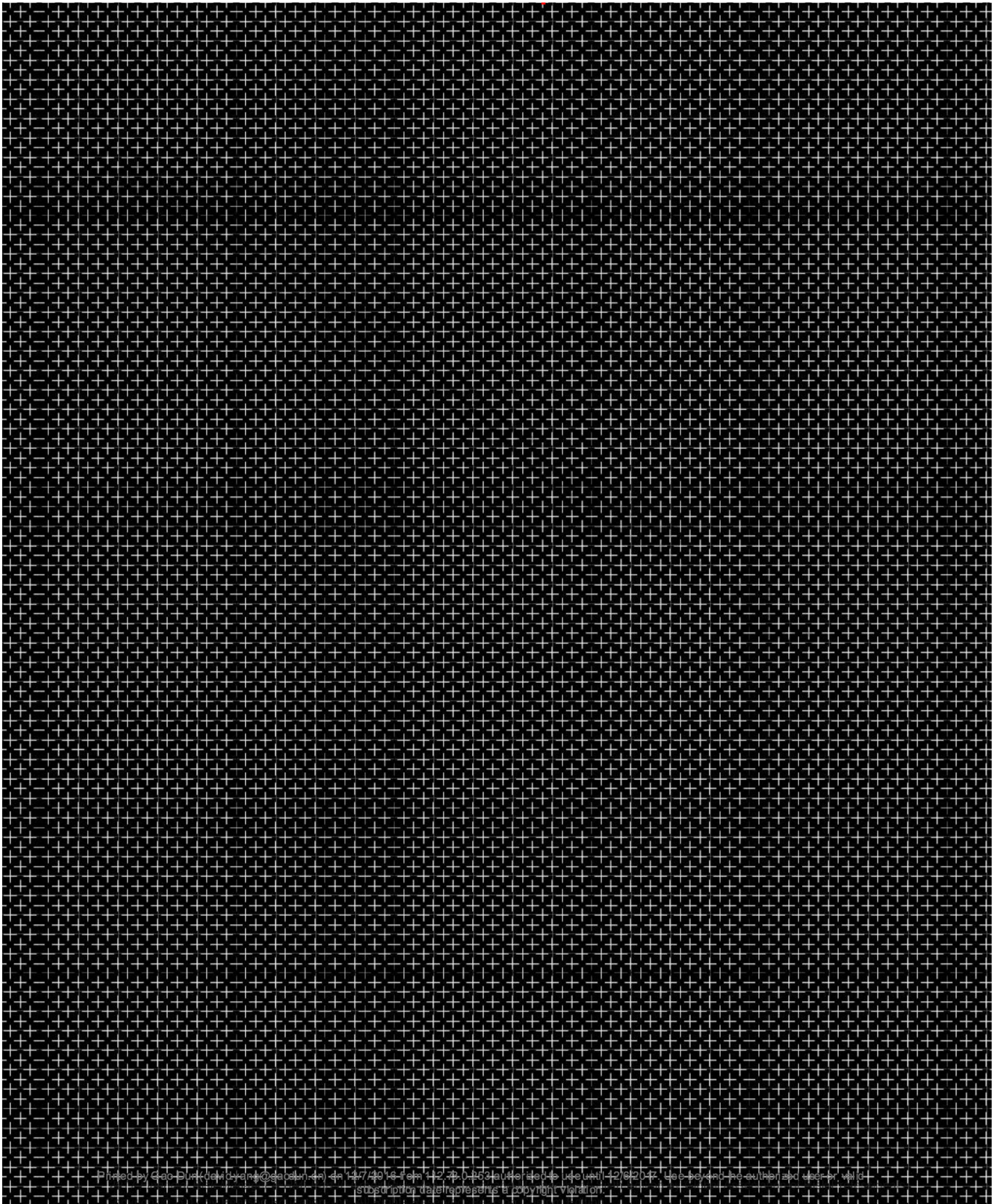
Duffie (1999), Hull and White (2000), and O'Kane and Turnbull (2003) provide overviews of CDS pricing.

Houweling and Vorst (2005) is an empirical study that finds hazard rate models to be reasonably accurate.

Klugman, Panjer, and Willmot (2008) provides an accessible introduction to hazard rate models. Litterman and Iben (1991); Berd, Mashal, and Wang (2003); and O'Kane and Sen (2004) apply hazard rate models to extract default probabilities. Schönbucher (2003), Chapters 4–5 and 7, is a clear exposition of the algebra.

See Markit Partners (2009) and Senior Supervisors Group (2009a) on the 2009 change in CDS conventions.





7

Portfolio Credit Risk

Learning Objectives

After completing this reading you should be able to:

- Define and calculate default correlation for credit portfolios.
- Identify drawbacks in using the correlation-based credit portfolio framework.
- Assess the impact of correlation on a credit portfolio and its Credit VaR.
- Describe the use of a single factor model to measure portfolio credit risk, including the impact of correlation.
- Define and calculate Credit VaR.
- Describe how Credit VaR can be calculated using a simulation of joint defaults.

Excerpt is Chapter 8 of Financial Risk Management: Models, History, and Institutions, *by Allan Malz.*

In this chapter, we extend the study of credit risk to portfolios containing several credit-risky securities. We begin by introducing the most important additional concept we need in this context, default correlation, and then discuss approaches to measuring portfolio credit risk.

A portfolio of credit-risky securities may contain bonds, commercial paper, off-balance-sheet exposures such as guarantees, as well as positions in credit derivatives such as credit default swaps (CDS). A typical portfolio may contain many different obligors, but may also contain exposures to different parts of one obligor's capital structure, such as preferred shares and senior debt. All of these distinctions can be of great importance in accurately measuring portfolio credit risk, even if the models we present here abstract from many of them.

In this chapter, we focus on an approach to measuring portfolio credit risk. It employs a factor model, the key feature of which is latent factors with normally distributed returns. Conditional on the values taken on by that set of factors, defaults are independent. There is a single future time horizon for the analysis. We will specialize the model even further to include only default events, and not credit migration, and only a single factor. In the CreditMetrics approach, this model is used to compute the distribution of credit migrations as well as default. One could therefore label the approach described in this chapter as "default-mode CreditMetrics." An advantage of this model is that factors can be related to real-world phenomena, such as equity prices, providing an empirical anchor for the model. The model is also tractable.

DEFAULT CORRELATION

In modeling a single credit-risky position, the elements of risk and return that we can take into consideration are

- The probability of default
- The loss given default (LGD), the complement of the value of recovery in the event of default
- The probability and severity of rating migration (nondefault credit deterioration)
- Spread risk, the risk of changes in market spreads for a given rating
- For distressed debt, the possibility of restructuring the firm's debt, either by negotiation among the owners of the firm and of its liabilities, or through the bankruptcy process. Restructuring opens the possibility of losses to owners of particular classes of debt as a result of a negotiated settlement or a judicial ruling

To understand credit portfolio risk, we introduce the additional concept of *default correlation*, which drives the likelihood of having multiple defaults in a portfolio of debt issued by several obligors. To focus on the issue of default correlation, we'll take default probabilities and recovery rates as given and ignore the other sources of return just listed.

Defining Default Correlation

The simplest framework for understanding default correlation is to think of

- Two firms (or countries, if we have positions in sovereign debt)
- With probabilities of default (or restructuring) π_1 and π_2
- Over some time horizon τ
- And a joint default probability—the probability that both default over τ—equal to π_{12}

This can be thought of as the distribution of the product of two Bernoulli-distributed random variables x_i, with four possible outcomes. We must, as in the single-firm case, be careful to define the Bernoulli trials as default or solvency over a specific time interval τ. In a portfolio credit model, that time interval is the same for all the credits in the book.

We have a new parameter π_{12} in addition to the single-name default probabilities. And it is a genuinely new parameter, a primitive: It is what it is, and isn't computed from π_1 and π_2, *unless* we specify it by positing that defaults are independent.

Since the value 1 corresponds to the occurrence of default, the product of the two Bernoulli variables equals 0 for three of the outcomes—those included in the event that at most one firm defaults—and 1 for the joint default event:

Outcome	x_1	x_2	x_1x_2	Probability
No default	0	0	0	$1 - \pi_1 - \pi_2 + \pi_{12}$
Firm 1 only defaults	1	0	0	$\pi_1 - \pi_{12}$
Firm 2 only defaults	0	1	0	$\pi_2 - \pi_{12}$
Both firms default	1	1	1	π_{12}

These are proper outcomes; they are distinct, and their probabilities add up to 1. The probability of the event that at least one firm defaults can be found as either 1 minus the probability of the first outcome, or the sum of the probabilities of the last three outcomes.

$$\mathbf{P}[\text{Firm 1 or Firm 2 or both default}] = \pi_1 + \pi_2 - \pi_{12}$$

We can compute the moments of the Bernoulli variates:

- The means of the two Bernoulli-distributed default processes are

$$\mathbf{E}[x_i] = \pi_i \qquad i = 1, 2$$

- The expected value of the product—representing joint default—is $\mathbf{E}[x_1 x_2] = \pi_{12}$.
- The variances are

$$\mathbf{E}[x_i]^2 - (\mathbf{E}[x_i])^2 = \pi_i(1 - \pi_i) \qquad i = 1, 2$$

- The covariance is

$$\mathbf{E}[x_1 x_2] - \mathbf{E}[x_1]\mathbf{E}[x_2] = \pi_{12} - \pi_1\pi_2$$

- The default correlation, finally, is

$$\rho_{12} = \frac{\pi_{12} - \pi_1\pi_2}{\sqrt{\pi_1(1-\pi_1)}\sqrt{\pi_2(1-\pi_2)}} \qquad \textbf{(7.1)}$$

We can treat the default correlation, rather than joint default probability, as the primitive parameter and use it to find the joint default probability:

$$\pi_{12} = \rho_{12}\sqrt{\pi_1(1-\pi_1)}\sqrt{\pi_2(1-\pi_2)} + \pi_1\pi_2$$

The joint default probability if the two default events are independent is $\pi_{12} = \pi_1\pi_2$, and the default correlation is $\rho_{12} = 0$. If $\rho_{12} \neq 0$, there is a linear relationship between the probability of joint default and the default correlation: The larger the "excess" of π_{12} over the joint default probability under independence, $\pi_1\pi_2$, the higher the correlation. Once we specify or estimate the π_i, we can nail down the joint default probability either directly or by specifying the default correlation. Most models, including those set out in this chapter, specify a default correlation rather than a joint default probability.

Example 7.1 Default Correlation

Consider a pair of credits, one BBB+ and the other BBB-rated, with $\pi_1 = 0.0025$ and $\pi_2 = 0.0125$. If the defaults are uncorrelated, then $\pi_{12} = 0.000031$, less than a third of a basis point. If, however, the default correlation is 5 percent, then $\pi_{12} = 0.000309$, nearly 10 times as great, and at 3 basis points, no longer negligible.

In a portfolio containing more than two credits, we have more than one joint default probability and default correlation. And, in contrast to the two-credit portfolio, we cannot specify the full distribution of defaults based just on the default probabilities and the pairwise correlations or joint default probabilities. To specify all the possible outcomes in a three-credit portfolio, we need the three single-default probabilities, the three two-default probabilities, and the no-default and three-default probabilities, a total of eight. But we have only seven conditions: the three single-default probabilities, three pairwise correlations, and the constraint that all the probabilities add up to unity. It's the latter constraint that ties out the probabilities when there are only two credits. With a number of credits $n > 2$, we have 2^n different events, but only $n + 1 + {}^{n(n-1)}\!/\!_2$ conditions:

n	2^n	$n + 1 + {}^{n(n-1)}\!/\!_2$
2	4	4
3	8	7
4	16	11
10	1,024	56

We can't therefore build an entire credit portfolio model solely on default correlations. But doing so is a pragmatic alternative to estimating or stipulating, say, the 1,024 probabilities required to fully specify the distribution of a portfolio of 10 credits.

Even if all the requisite parameters could be identified, the number would be quite large, since we would have to define a potentially large number of pairwise correlations. If there are N credits in the portfolio, we need to define N default probabilities and N recovery rates. In addition, we require $N(N-1)$ pairwise correlations. In modeling credit risk, we often set all of the pairwise correlations equal to a single parameter. But that parameter must then be non-negative, in order to avoid correlation matrices that are not positive-definite and results that make no sense: Not all the firms' events of default can be negatively correlated with one another.

Example 7.2

Consider a portfolio containing five positions:

1. A five-year senior secured bond issued by Ford Motor Company

2. A five-year subordinate unsecured bond issued by Ford Motor Company
3. Long protection in a five-year CDS on Ford Motor Credit Company
4. A five-year senior bond issued by General Motors Company
5. A 10-year syndicated term loan to Starwood Resorts

If we set a horizon for measuring credit risk of $\tau = 1$ year, we need to have four default probabilities and 12 pairwise default correlations, since there are only four distinct corporate entities represented in the portfolio. However, since the two Ford Motor Company bonds are at two different places in the capital structure, they will have two different recovery rates.

This example has omitted certain types of positions that will certainly often occur in real-world portfolios. Some of their features don't fit well into the portfolio credit risk framework we are developing:

- Guarantees, revolving credit agreements, and other contingent liabilities behave much like credit options.
- CDS basis trades are not essentially market- or credit-risk–oriented, although both market and credit risk play a very important role in their profitability. Rather, they may be driven by "technical factors," that is, transitory disruptions in the typical positioning of various market participants.

  A dramatic example occurred during the subprime crisis. The CDS basis widened sharply as a result of the dire lack of funding liquidity.
- Convertible bonds are both market- and credit-risk oriented. Equity and equity vega risk can be as important in convertible bond portfolios as credit risk.

The Order of Magnitude of Default Correlation

For most companies that issue debt, most of the time, default is a relatively rare event. This has two important implications:

1. Default correlation is hard to measure or estimate using historical default data. Most studies have arrived at one-year correlations on the order of 0.05. However, estimated correlations vary widely for different time periods, industry groups, and domiciles, and are often negative.
2. Default correlations are small in magnitude.

In other contexts, for example, thinking about whether a regression result indicates that a particular explanatory value is important, we get used to thinking of, say, 0.05 as a "small" or insignificant correlation and 0.5 as a large or significant one. The situation is different for default correlations because probabilities of default tend to be small—on the order of 1 percent—for all but the handful of CCC and below firms. The probability of any particular pair of credits defaulting is therefore also small, so an "optically" small correlation can have a large impact, as we saw in Example 7.1.

CREDIT PORTFOLIO RISK MEASUREMENT

To measure credit portfolio risk, we need to model default, default correlation, and loss given default. In more elaborate models, we can also include ratings migration. We restrict ourselves here to default mode. But in practice, and in such commercial models as Moody's KMV and CreditMetrics, models operate in migration mode; that is, credit migrations as well as default can occur.

Granularity and Portfolio Credit Value-at-Risk

Portfolio Credit VaR is defined similarly to the VaR of a single credit. It is a quantile of the credit loss, minus the expected loss of the portfolio.

Default correlation has a tremendous impact on portfolio risk. But it affects the volatility and extreme quantiles of loss rather than the expected loss. If default correlation in a portfolio of credits is equal to 1, then the portfolio behaves as if it consisted of just one credit. No credit diversification is achieved. If default correlation is equal to 0, then the number of defaults in the portfolio is a binomially distributed random variable. Significant credit diversification may be achieved.

To see how this works, let's look at diversified and undiversified portfolios, at the two extremes of default correlation, 0 and 1. Imagine a portfolio of n credits, each with a default probability of π percent and a recovery rate of zero percent. Let the total value of the portfolio be $1,000,000,000. We will set n to different values, thus dividing the portfolio into larger or smaller individual positions. If $n = 50$, say, each position has a value of

$20,000,000. Next, assume each credit is in the same place in the capital structure and that the recovery rate is zero; in the event of default, the position is wiped out. We'll assume each position is an obligation of a different obligor; if two positions were debts of the same obligor, they would be equivalent to one large position. We can either ignore the time value of money, which won't play a role in the example, or think of all of these quantities as future values.

Now we'll set the default correlation to either 0 or 1.

- If the default correlation is equal to 1, then either the entire portfolio defaults, with a probability of π, or none of the portfolio defaults. In other words, with a default correlation of 1, regardless of the value of n, the portfolio behaves as though $n = 1$.

  We can therefore continue the analysis by assuming all of the portfolio is invested in one credit. The expected loss is equal to $\pi \times 1{,}000{,}000{,}000$. But with only one credit, there are only the two all-or-nothing outcomes. The credit loss is equal to 0 with probability $1 - \pi$. The default correlation doesn't matter.

  The extreme loss given default is equal to $1,000,000,000, since we've assumed recovery is zero. If π is greater than the confidence level of the Credit VaR, then the VaR is equal to the entire $1,000,000,000, less the expected loss. If π is less than the confidence level, then the VaR is less than zero, because we always subtract the expected from the extreme loss. If, for example, the default probability is $\pi = 0.02$, the Credit VaR at a confidence level of 95 percent is negative (i.e., a gain), since there is a 98 percent probability that the credit loss in the portfolio will be zero. Subtracting from that the expected loss of $\pi \times 1{,}000{,}000{,}000 = 20{,}000{,}000$ gives us a VaR of –$20,000,000. The Credit VaR in the case of a single credit with binary risk is well-defined and can be computed, but not terribly informative.

- If the default correlation is equal to 0, the number of defaults is binomially distributed with parameters n and π. We then have many intermediate outcomes between the all-or-nothing extremes.

  Suppose there are 50 credits in the portfolio, so each position has a future value, if it doesn't default, of $20,000,000. The expected loss is the same as with one credit: $\pi \times 1{,}000{,}000{,}000$. But now the extreme outcomes are less extreme. Suppose again that $\pi = 0.02$. The number of defaults is then binomially distributed with parameters 50 and 0.02. The 95th percentile of the number of defaults is 3, as seen in Figure 7-1; the probability of two defaults or less is 0.92 and the probability of three defaults or less is 0.98. With three defaults, the credit loss is $60,000,000. Subtracting the expected loss of $20,000,000, which is the same as for the single-credit portfolio, we get a Credit VaR of $40,000,000.

  As we continue to increase the number of positions and decrease their size, keeping the total value of the portfolio constant, we decrease the variance of portfolio values. For $n = 1{,}000$, the 95th percentile of defaults is 28, and the 95th percentile of credit loss is $28,000,000, so the Credit VaR is $8,000,000.

We summarize the results for $n = 1$, 50, 1,000, for default probabilities $\pi = 0.005$, 0.02, 0.05, and at confidence levels of 95 and 99 percent in Table 7-1 and in Figure 7-2.

What is happening as the portfolio becomes more *granular*, that is, contains more independent credits, each of

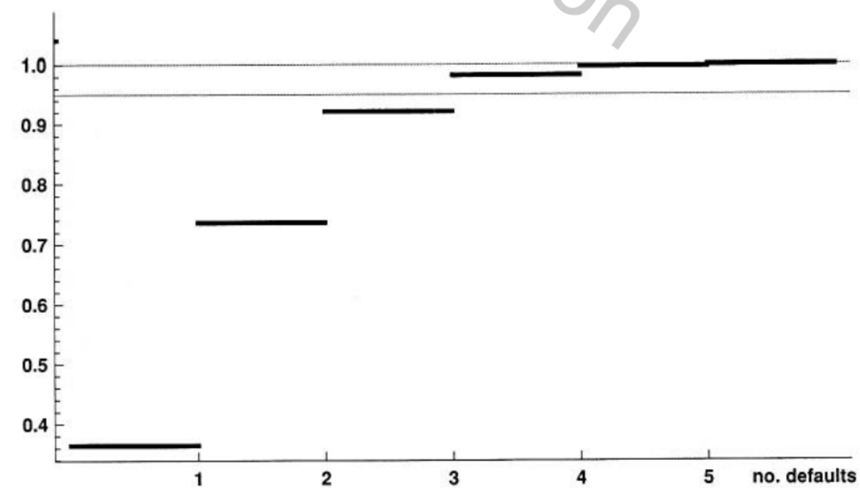


FIGURE 7-1 Distribution of defaults in an uncorrelated credit portfolio cumulative probability distribution function of the number of defaults in a portfolio of 50 independent credits with a default probability of 2 percent.

TABLE 7-1 Credit VaR of an Uncorrelated Credit Portfolio

	$\pi = 0.005$	$\pi = 0.02$	$\pi = 0.05$
Expected loss	5,000,000	20,000,000	50,000,000
	$n = 1$		
	95 Percent Confidence Level		
Number of defaults	0	1	1
Proportion of defaults	0.000	0.000	0.000
Credit value-at-risk	−5,000,000	−20,000,000	−50,000,000
	99 Percent Confidence Level		
Number of defaults	0	1	1
Proportion of defaults	0.000	1.000	1.000
Credit value-at-risk	−5,000,000	980,000,000	950,000,000
	$n = 50$		
	95 Percent Confidence Level		
Number of defaults	1	3	5
Proportion of defaults	0.020	0.060	0.100
Credit value-at-risk	15,000,000	40,000,000	50,000,000
	99 Percent Confidence Level		
Number of defaults	2	4	7
Proportion of defaults	0.040	0.080	0.140
Credit value-at-risk	35,000,000	60,000,000	90,000,000
	$n = 1000$		
	95 Percent Confidence Level		
Number of defaults	9	28	62
Proportion of defaults	0.009	0.028	0.062
Credit value-at-risk	4,000,000	8,000,000	12,000,000
	99 Percent Confidence Level		
Number of defaults	11	31	67
Proportion of defaults	0.011	0.031	0.067
Credit value-at-risk	6,000,000	11,000,000	17,000,000

which is a smaller fraction of the portfolio? The Credit VaR is, naturally, higher for a higher probability of default, given the portfolio size. But it decreases as the credit portfolio becomes more granular for a given default probability. The convergence is more drastic with a high default probability. But that has an important converse: It is harder to reduce VaR by making the portfolio more granular, if the default probability is low.

Eventually, for a credit portfolio containing a very large number of independent small positions, the probability converges to 100 percent that the credit loss will equal the expected loss. While the single-credit portfolio experiences no loss with probability $1 - \pi$ and a total loss with probability π, the granular portfolio experiences a loss of 100π percent "almost certainly." The portfolio then has zero volatility of credit loss, and the Credit VaR is zero.

In the rest of this chapter, we show how models of portfolio credit risk take default correlation into account, focusing on one model in particular: The single-factor model, since it is a structural model, emphasizes the correlation between the fundamental driver of default of different firms. Default correlation in that model depends on how closely firms are tied to the broader economy.

DEFAULT DISTRIBUTIONS AND CREDIT VaR WITH THE SINGLE-FACTOR MODEL

In the example of the last section, we set default correlation only to the extreme values of 0 and 1, and did not take account of idiosyncratic credit risk. In the rest of this chapter, we permit default correlation to take values anywhere on (0, 1). The single-factor model enables us to vary default correlation through the credit's beta to the market factor and lets idiosyncratic risk play a role.

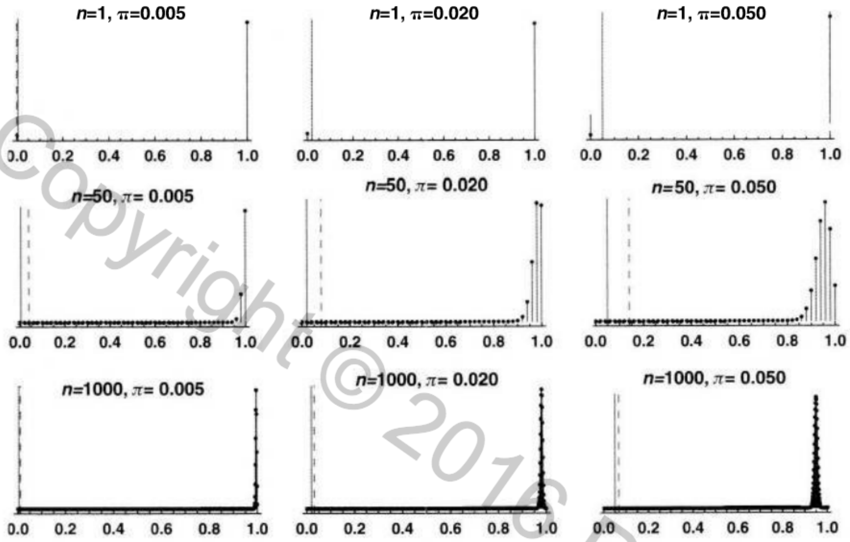


FIGURE 7-2 Distribution of losses in an uncorrelated credit portfolio.

The graph displays the probability density of losses for each combination of a number of equally sized credits and default probabilities. The initial future value of the portfolio is $1,000,000,000. The values on the x-axis can be interpreted as the fraction of credit losses or as the dollar value of loss in billions. The dashed grid line marks the 99th percentile of loss. The solid grid line marks the expected loss and is the same in each panel.

Conditional Default Distributions

To use the single-factor model to measure portfolio credit risk, we start by imagining a number of firms $i = 1, 2, \ldots$, each with its own correlation β_i to the market factor, its own standard deviation of idiosyncratic risk $\sqrt{1-\beta_i^2}$, and its own idiosyncratic shock ϵ_i. Firm i's return on assets is

$$a_i = \beta_i m + \sqrt{1-\beta_i^2}\epsilon_i \qquad i = 1, 2, \ldots$$

We assume that m and ϵ_i are standard normal variates, and are not correlated with one another. We now in addition assume the ϵ_i are not correlated with one another:

$$m \sim N(0, 1)$$

$$\epsilon_i \sim N(0, 1) \quad i = 1, 2, \ldots$$

$$\mathrm{Cov}[m, \epsilon_i] = 0 \quad i = 1, 2, \ldots$$

$$\mathrm{Cov}[\epsilon_i, \epsilon_j] = 0 \quad i, j = 1, 2, \ldots$$

Under these assumptions, each a_i is a standard normal variate. Since both the market factor and the idiosyncratic shocks are assumed to have unit variance, the beta of each credit i to the market factor is equal to β_i. The correlation between the asset returns of any pair of firms i and j is $\beta_i\beta_j$:

$$\mathbf{E}[a_i] = 0 \quad i = 1, 2, \ldots$$

$$\mathrm{var}[a_i] = \beta_i^2 + 1 - \beta_i^2 = 1 \quad i = 1, 2, \ldots$$

$$\mathrm{Cov}[a_i, a_j] = E\left[\left(\beta_i m + \sqrt{1-\beta_i^2}\epsilon_i\right)\left(\beta_j m + \sqrt{1-\beta_j^2}\epsilon_j\right)\right]$$

$$= \beta_i\beta_j \; i, j = 1, 2, \ldots$$

Just as in the single-credit version of the model, firm i defaults if $a_i \leq k_i$, the logarithmic distance to the default asset value, measured in standard deviations.

Example 7.3 Correlation and Beta in Credit Single-Factor Model

Suppose firm 1 is "cyclical" and has $\beta_1 = 0.5$, while firm 2 is "defensive" and has $\beta_2 = 0.1$. The asset return correlation of the two firms is then $\beta_1\beta_2 = 0.5 \times 0.1 = 0.05$.

The single-factor model has a feature that makes it an especially handy way to estimate portfolio credit risk: conditional independence, the property that once a particular value of the market factor is realized, the asset returns—and hence default risks—are independent of one another. Conditional independence is a result of the model assumption that the firms' returns are correlated only via their relationship to the market factor.

To see this, let m take on a particular value $\bar{m}$. The distance to default—the asset return—increases or decreases, and now has only one random driver ϵ_i, the idiosyncratic shock:

$$a_i - \beta_i \bar{m}\sqrt{1-\beta_i^2}\epsilon_i \qquad i = 1, 2, \ldots$$

The mean of the default distribution shifts for any $\beta_i > 0$ when the market factor takes on a specific value. The variance of the default distribution is reduced from 1 to $\sqrt{1-\beta_i^2}$, even though the default threshold k_i has not changed. The change in the distribution that results from conditioning is illustrated in Figure 7-3.

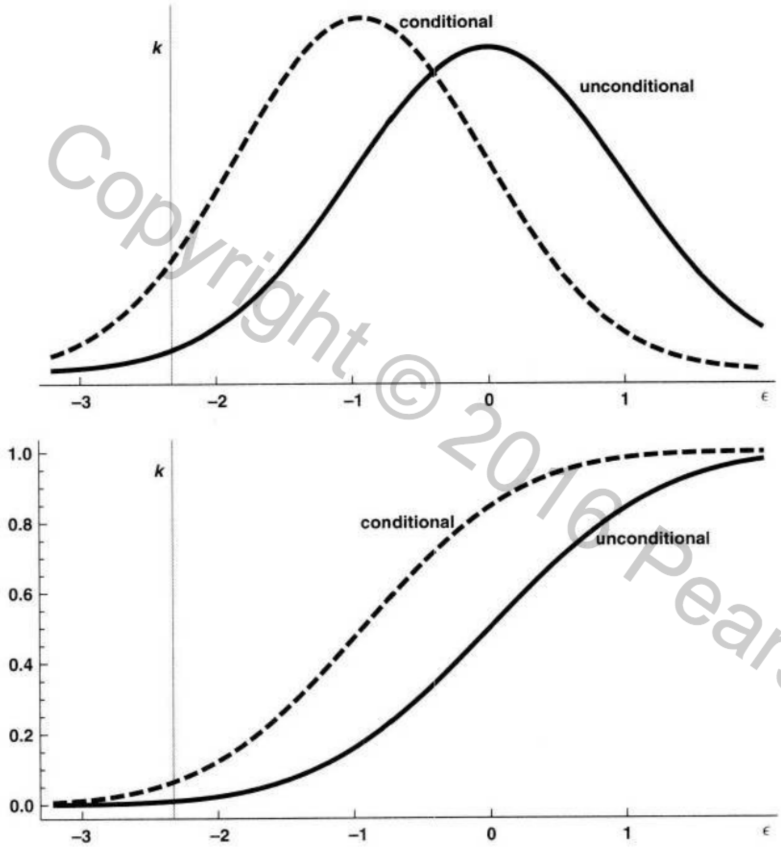


FIGURE 7-3 Default probabilities in the single-factor model.

The graph assumes $\beta_i = 0.4$, $k_i = -2.33$ ($\Leftrightarrow \pi_i = 0.01$), and $\bar{m} = -2.33$. The unconditional default distribution is a standard normal distribution, while the conditional default distribution is $N\left(\beta_i \bar{m}, \sqrt{(1-\beta_i^2)}\right)$, $= N(-0.4, 0.9165)$.
Upper panel: Unconditional and conditional probability density of default. Note that the mean as well as the volatility of the conditional distribution are lower.
Lower panel: Unconditional and conditional cumulative default distribution function.

Example 7.4 Default Probability and Default Threshold

Suppose a firm has $\beta_i = 0.4$ and k_i −2.33, it is a middling credit, but cyclical (relatively high β_i). Its unconditional probability of default is $\Phi(-2.33) = 0.01$. If we enter a modest economic downturn, with $\bar{m} = -1.0$, the conditional asset return distribution is $N(-0.4, \sqrt{1-0.40^2})$ or $N(-0.4, 0.9165)$, and the conditional default probability is found by computing the probability that this distribution takes on the value −2.33. That probability is 1.78 percent.

If we were in a stable economy with $m = 0$, we would need a shock of −2.33 standard deviations for the firm to die. But with the firm's return already 0.4 in the hole because of an economy-wide recession, it takes only a 1.93 standard deviation *additional* shock to kill it.

Now suppose we have a more severe economic downturn, with $\bar{m} = -2.33$. The firm's conditional asset return distribution is $N(-0.932, 0.9165)$ and the conditional default probability is 6.4 percent. A 0.93 standard deviation shock ($\epsilon_i \leq -0.93$) will now trigger default.

To summarize, specifying a realization $m = \bar{m}$ does three things:

1. The conditional probability of default is greater or smaller than the unconditional probability of default, unless either $\bar{m} = 0$ or $\beta_i = 0$, that is, either the market factor shock happens to be zero, or the firm's returns are independent of the state of the economy.

   There is also no longer an infinite number of combinations of market and idiosyncratic shocks that would trigger a firm i default. Given $\bar{m}$, a realization of ϵ_i less than or equal to

   $$k_i - \beta_i \bar{m} \qquad i = 1, 2, \ldots$$

   triggers default. This expression is linear and downward sloping in $\bar{m}$: As we let $\bar{m}$ vary from high (strong economy) to low (weak economy) values, a smaller (less negative) idiosyncratic shock will suffice to trigger default.
2. The conditional variance of the default distribution is $1 - \beta_i^2$, so the conditional variance is reduced from the unconditional variance of 1.
3. It makes the asset returns of different firms *independent.* The ϵ_i are independent, so the conditional returns $\sqrt{1-\beta_i^2}\epsilon_i$ and $\sqrt{1-\beta_j^2}\epsilon_j$ and thus the default outcomes for two different firms i and j are independent.

Putting this all together, while the unconditional default distribution is a standard normal, the conditional distribution can be represented as a normal with a mean of $-\beta_i \bar{m}$ and a standard deviation of $\sqrt{1-\beta_i^2}$.

The *conditional cumulative default probability function* can now be represented as a function of m:

$$p(m) = \Phi\left(\frac{k_i - \beta_i m}{\sqrt{1-\beta_i^2}}\right) \qquad i = 1, 2, \ldots$$

It is plotted in the lower panel of Figure 7-4 for different correlations. This function is the standard normal distribution function of a random variable that has been standardized in a specific way. The mean, or "number of standard deviations," is set to the new distance to default, given the realization of the market factor, while the standard deviation itself is set to its value $\sqrt{1-\beta_i^2}$ under conditional independence. The intuition is that, for a given value of the market factor, the probability of default depends on how many standard deviations below its mean of 0 is the realization of ϵ_i. The density function corresponding to the cumulative default function is plotted in the upper panel of Figure 7-4.

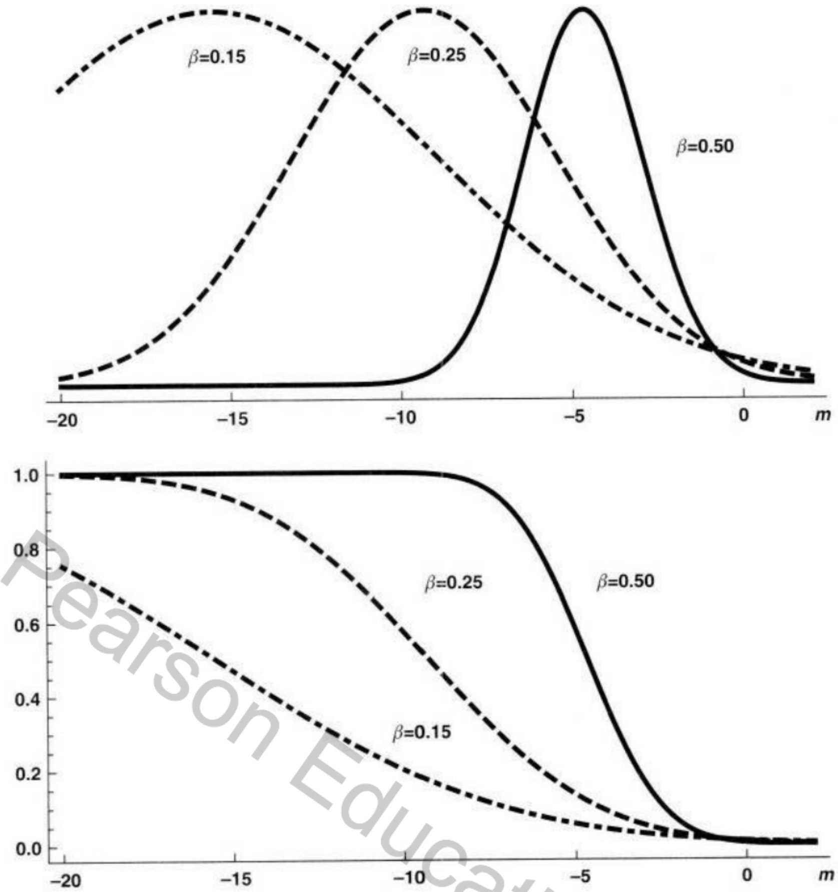


FIGURE 7-4 Single-factor default probability distribution.

Probability of default of a single obligor, conditional on the realization of m (x-axis). The default probability is set to 1 percent ($k = -2.33$), and the default correlation is set to different values as specified by the plot labels.
Upper panel: Conditional default density function, that is, the density function corresponding to $p(m)$.
Lower panel: Conditional cumulative distribution function of default $p(m)$.

Asset and Default Correlation

We began earlier to discuss the difference between the asset return and the default correlation. Let's look for a moment at the relationship between the two.

In the single-factor model, the cumulative return distribution of any pair of credits i and j is a bivariate standard normal with a correlation coefficient equal to $\beta_i\beta_j$:

$$\begin{pmatrix} a_i \\ a_j \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & \beta_i\beta_j \\ \beta_i\beta_j & 1 \end{pmatrix}\right]$$

Its cumulative distribution function is $\Phi\binom{a_i}{a_j}$. The probability of a joint default is then equal to the probability that the realized value is in the region $\{-\infty \le a_i \le k_i, -\infty \le a_j \le k_j\}$:

$$\Phi\begin{pmatrix} k_i \\ k_j \end{pmatrix} = \mathbf{P}[-\infty \le a_i \le k_i,\ -\infty \le a_j \le k_j]$$

To get the default correlation for this model, we substitute $\pi_{ij} = \Phi\binom{k_i}{k_j}$ into Equation (7.1), the expression for the linear correlation:

$$\rho_{ij} = \frac{\Phi\binom{k_i}{k_j} - \pi_i\pi_j}{\sqrt{\pi_i(1-\pi_i)}\sqrt{\pi_j(1-\pi_j)}}$$

From here on, let's assume that the parameters are the same for all firms; that is, $\beta_i = \beta$, $k_i = k$, and $\pi_i = \pi$, $i = 1, 2, \ldots$ The pairwise asset return correlation for any two firms is then β^2. The probability of a joint default for any two firms for this model is

$$\Phi\begin{pmatrix} k \\ k \end{pmatrix} = \mathbf{P}\,[-\infty \le a \le k, -\infty \le a \le k]$$

and the default correlation between any pair of firms is

$$\rho = \frac{\Phi\binom{k}{k} - \pi^2}{\pi(1-\pi)}$$

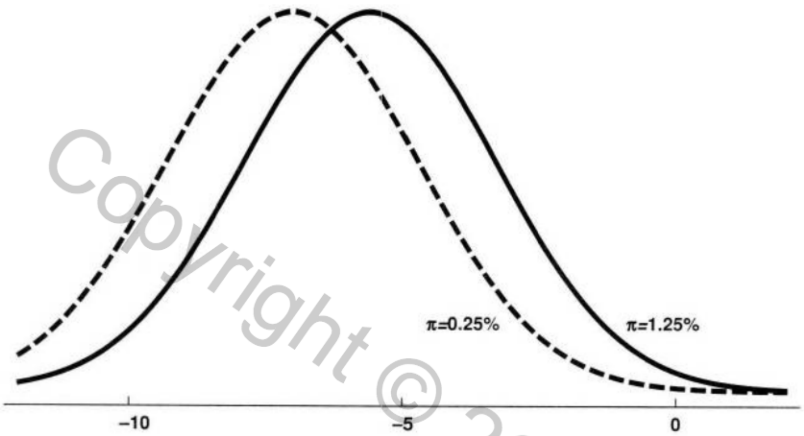


FIGURE 7-5 Conditional default density function in the single-factor model.

Both plots take $\beta = 0.40$. For a given correlation, the probability of default changes the location of the default distribution, but not its variance.

Example 7.5 Default Correlation and Beta

What β corresponds to a "typical" low investment-grade default probability of 0.01 and a default correlation of 0.05? We need to use a numerical procedure to find the parameter β that solves

$$\rho = 0.05 = \frac{\Phi\binom{k}{k} - \pi^2}{\pi(1-\pi)}$$

With $\pi = 0.01$, the results are $\beta = 0.561$, the asset correlation $\beta^2 = 0.315$, and a joint default probability of 0.0006, or 6 basis points. Similarly, starting with $\beta = 0.50$ ($\beta^2 = 0.25$), we find a joint default probability of 4.3 basis points and a default correlation of 0.034.

Credit VaR Using the Single-Factor Model

In this section, we show how to use the single-factor model to estimate the Credit VaR of a "granular," homogeneous portfolio. Let n represent the number of firms in the portfolio, and assume n is a large number. We will assume the loss given default is \$1 for each of the n firms. Each credit is only a small fraction of the portfolio and idiosyncratic risk is de minimis.

Conditional Default Probability and Loss Level

Recall that, for a given realization of the market factor, the asset returns of the various credits are independent standard normals. That, in turn, means that we can apply the law of large numbers to the portfolio. For each level of the market factor, the *loss level* $x(m)$, that is, the fraction of the portfolio that defaults, converges to the conditional probability that a single credit defaults, given for any credit by

$$p(m) = \Phi\left(\frac{k - \beta m}{\sqrt{1-\beta^2}}\right) \qquad \textbf{(7.2)}$$

So we have

$$\lim_{N \to \infty} x(m) = p(m) \quad \forall m \in \mathbb{R}$$

The intuition is that, if we know the realization of the market factor return, we know the level of losses realized. This in turn means that, given the model's two parameters, the default probability and correlation, portfolio returns are driven by the market factor.

Unconditional Default Probability and Loss Level

We are ultimately interested in the unconditional, not the conditional, distribution of credit losses. The unconditional probability of a particular loss level is equal to the probability that the the market factor return that leads to that loss level is realized. The procedure for finding the unconditional distribution is thus:

1. Treat the loss level as a random variable X with realizations x. We don't simulate x, but rather work through the model analytically for each value of x between 0 (no loss) and 1 (total loss).
2. For each level of loss x, find the realization of the market factor at which, for a single credit, default has a probability equal to the stated loss level. The loss level and the market factor return are related by

$$x(m) = p(m) = \Phi\left(\frac{k - \beta m}{\sqrt{1-\beta^2}}\right)$$

So we can solve for $\bar{m}$, the market factor return corresponding to a given loss level $\bar{x}$:

$$\Phi^{-1}(\bar{x}) = \frac{k - \beta\bar{m}}{\sqrt{1-\beta^2}}$$

or

$$\bar{m} = \frac{k - \sqrt{1-\beta^2}\Phi^{-1}(\bar{x})}{\beta}$$

3. The probability of the loss level is equal to the probability of this market factor return. But by assumption, the market factor is a standard normal:

$$\mathbf{P}[X \leq \bar{x}] = \Phi(\bar{m}) = \Phi\left(\frac{k - \sqrt{1-\beta^2}\Phi^{-1}(\bar{x})}{\beta}\right)$$

4. Repeat this procedure for each loss level to obtain the probability distribution of X.

Another way of describing this procedure is: Set a loss level/conditional default probability x and solve the conditional cumulative default probability function, Equation (7.2), for $\bar{m}$ such that:

$$\bar{m} = \frac{k - \sqrt{1-\beta^2}\Phi^{-1}(x)}{\beta}$$

The loss distribution function is thus

$$\mathbf{P}[X \leq x] = \Phi\left(\frac{k - \sqrt{1-\beta^2}\Phi^{-1}(x)}{\beta}\right)$$

Example 7.6 Loss Level and Market Level

A loss of 0.01 or worse occurs when—converges to the event that—the argument of $p(m)$ is at or below the value such that $p(m) = 0.01$.

$$p(\bar{m}) = 0.01 = \Phi\left(\frac{k - \beta\bar{m}}{\sqrt{1-\beta^2}}\right)$$

The value $\bar{m}$ at which this occurs is found by solving

$$\Phi^{-1}(0.01) \approx -2.33 = p^{-1}(\bar{m}) = \frac{k - \beta\bar{m}}{\sqrt{1-\beta^2}}$$

for $\bar{m}$. This is nothing more than solving for the $\bar{m}$ that gives you a specific quantile of the standard normal distribution.

With a default probability $\pi = 0.01$ and correlation $\beta^2 = 0.50^2 = 0.25$, the solution is $\bar{m} = -0.6233$. The probability that the market factor ends up at -0.6233 or less is $\Phi(-0.6233) = 0.2665$.

As simple as the model is, we have several parameters to work with:

- The probability of default π sets the unconditional expected value of defaults in the portfolio.
- The correlation to the market β^2 determines how spread out the defaults are over the range of the market factor. When the correlation is high, then, for any probability of default, defaults mount rapidly as business conditions deteriorate. When the correlation is low, it takes an extremely bad economic scenario to push the probability of default high.

To understand the impact of the correlation parameter, start with the extreme cases:

- $\beta \rightarrow 1$ (perfect correlation). Recall that we have constructed a portfolio with no idiosyncratic risk. If the correlation to the market factor is close to unity, there are two possible outcomes. Either $m \leq k$, in which case nearly all the credits default, and the loss rate is equal to 1, or $m > k$, in which case almost none default, and the loss rate is equal to 0.
- $\beta \rightarrow 0$ (zero correlation). If there is no statistical relationship to the market factor, so idiosyncratic risk is nil, then the loss rate will very likely be very close to the default probability p.

In less extreme cases, a higher correlation leads to a higher probability of either very few or very many defaults, and a lower probability of intermediate outcomes.

Further Reading

Lucas (1995) provides a definition of default correlation and an overview of its role in credit models. See also Hull and White (2001).

Credit Suisse First Boston (2004) and Lehman Brothers (2003) are introductions by practitioners. Zhou (2001) presents an approach to modeling correlated defaults based on the Merton firm value, rather than the factor-model approach.

The application of the single-factor model to credit portfolios is laid out in Finger (1999) and Vasicek (1991). Accessible introduction to copula theory are Frees and Valdez (1998) and in Klugman, Panjer, and Willmot (2008). The application to credit portfolio models and the equivalence to Gaussian CreditMetrics is presented in Li (2000).

The correlated intensities approach to modeling credit portfolio risk, as well as other alternatives to the Gaussian single-factor approach presented here, are described in Schönbucher (2003), Chapter 10, and Lando (2004), Chapter 5.