Credit risk modelling

Best practice in credit risk modelling and management combines empirical data, research expertise, and technological capability. In this article, Geoff Fite and Jing Zhang identify and expound on these requirements and illustrate a sample solution that incorporates them

The state of credit risk measurement has been evolving rapidly since the last credit cycle. Best practice in credit risk modelling and management is now, more than ever and irrespective of portfolio size and institution characteristics, dependent upon three critical capabilities: empirical data, research expertise and technological capability. Successful management of credit risk begins with the measurement of individual obligors and instruments, culminating in the analysis of complex portfolios of varied exposures.

Empirical data

The importance of empirical data to quantitative credit risk measurement cannot be overemphasised. There are key three requirements of empirical data in its effective use for risk modelling and measurement. The first requirement is that specific information must be collected on each individual obligor and risk must first be measured at this level. This includes counterparty identifiers, financial statements, country and industry classifications and obligor relationship information, such as parent and subsidiary, guarantor and guarantee, etc. From these inputs, estimates of default and migration risk can be quantitatively measured in a consistent fashion across all obligors. Regional and industry variations in data characteristics must be accounted for and country or industry-specific models should be tailored to accommodate data validity issues while, at the same time, driving towards consistent measurement of individual obligor risk globally.

The second data requirement is at individual exposure level. This includes origination data such as exposure amount, pricing information and terms and conditions for loans and bonds. Post-origination information of drawn amount, payment histories, default and recovery information needs to be tracked and linked to the obligor information, such as obligor risk migration, support and guarantee information by parent and grantors, etc. For derivatives, this includes market prices, structures of the contracts and their exposure profiles over time.

The third and most challenging requirement, from the perspective of building quantitative credit risk models, is that this information must have sufficient breadth and depth. The nature of credit events such as defaults and ultimate recoveries is that they are often infrequent and opaque. Under the impetus of Basel II and other regulatory requirements, the majority of financial institutions have only recently started collecting default and recovery information. Furthermore, best practice dictates a correlation model that accounts for the linking and interdependence of the global, regional, country and industry economic factors. This implies collection of obligor and instrument information, and their ultimate performance, at a scale well beyond the possibility of any single credit market participant. More often than not, it is implausible that individual institutions can build credit risk models based solely on the historical data in their individual portfolios. The solution is to pool and share data by forming consortia of peer institutions and/or to seek help from firms that specialise in collecting and standardising such information over time.

Research expertise

Once a foundation of empirical data is laid, it is then possible to build successful credit risk models. Constructing models is a thoughtful and laborious process of balancing and weaving business objectives, economic theories, quantitative techniques and empirical facts. Naturally, business objectives should take precedence as any model is only useful to the extent it is applied in practice. The real test of any risk model is whether it is intuitive, adds value to the management of risk at both the individual instrument and portfolio levels and can be reasonably calculated given the required inputs. A very accurate model with hundreds of variables is not very useful to institutions that cannot populate all of the inputs, and will not gain acceptance if it is unintuitive and overly complex. The guidance of sound economic theory helps build models that are intuitive, easy to explain and are robust across counterparties and instruments over time. Building a model is not simply about producing the most accurate numbers (although this is essential), but also about opening up the process for determining credit risk in a way that is useful to the practitioner.

Furthermore, risk models must accommodate all possible instruments in order to accurately and comprehensively measure the correlation and concentration effect within potentially complex portfolios. The required research expertise for successfully building these models is very much cross-functional in nature, ranging from financial economics, accounting and statistics to operational research and mathematics. The synergy of such a cross-function effort is typically manifested in the innovation and sophistication embedded in these models. Research of credit risk measurement requires data with sufficient breadth and depth, modern statistical skills and intimate knowledge of data. The ultimate test of the success of any quantitative model is whether it can be empirically shown to work in real business practice.

Technological capability

The breadth and amount of empirical data required to construct, maintain and implement today's best credit risk models represent a technology challenge of the highest order. The technological systems that measure and manage credit risk require expertise in data management, software construction and physical architecture design.

Data management is a process that involves the collection and processing of data over time in a consistent manner for all obligors, instruments and events. Snapshots of data are generally insufficient to drive empirical data requirements, so best practice in data management is to think of data as the product of a refinery. Raw information, from multiple and sometimes overlapping sources, is brought into the front of the refinery. The first step is to cleanse and standardise the raw data into consistent formats and to cross-check information across multiple sources. Human involvement is inevitable in the cleansing and standardisation process but, because the refinery must be running 24 hours a day, seven days a week, it is necessary to have a follow-the-sun data-cleansing operation. Clean and standardised data is then run through quantitative models and, in some cases, qualitative processes to measure the risk of individual obligors and instruments. Information on credit information such as default, migration and recovery should also be processed in the refinery and linked to other related data items, so that ultimately a credit risk models can be estimated across inputs that are appropriate for the asset class and exposure. Finally, data management of credit portfolios is a challenge of taking the outputs from a refinery that calculates obligor and instrument risk and combining them with exposures held at a financial institution. A well-run data refinery of global data will contain upwards of 50 terabytes of information, and financial institution portfolios can also be extremely large; as a result, the data management process should be a core competence of the group performing risk measurement and implementing portfolio management solutions.

Software construction in credit risk modelling is no less complex than the data management required to feed it. The best software solutions incorporate independently validated quantitative models, scalable application architecture, role-based security, integration points with both internal and external systems, intuitive user interfaces and high-performance computing infrastructure. Best practice in model development is to develop model prototypes in statistical and mathematical packages such as SAS or MatLab as a means to test and calibrate mathematical models using a variety of empirical data sets.

Once the prototype is validated, the model should be converted to a high-performance runtime language such as C++ or Java since compiled statistical package code generally does not perform at the speed required for large calculation runs. The engine code that contains all models should be scalable to accommodate multiple concurrent methodologies, yet – at the same time – insulate the calculations from the rest of the application logic to ensure system stability. Role-based security covering the common use cases at a financial institution is required to allow multiple business requirements to be met using the same software platform. For example, one group at a financial institution may be engaged in spreading financials of individual firms, another group may be conducting origination and setting price using *ex ante* deal analysis, another group will be running capital adequacy calculations across an individual portfolio and another may be providing institution-wide reporting for regulators and executive management.

Data integration points are required to support data feeds from third parties for public company information, instrument pricing and supporting data, and to also link loan accounting, trading and limits management systems. User interfaces should be intuitive and must be able to support local languages and multiple locations concurrently. Finally, the nature of large data sets, computationally intensive calculations across them and the need to support multiple roles leads to the requirement that the physical infrastructure include software as a service, grid computing and high-performance database technology.

Sample solution

A typical implementation of best practice credit risk measurement systems illustrates these requirements. For example, a large multinational bank may have multiple member banks around the world, with a complex portfolio of exposures to private firms, public companies, retail consumers, commercial real estate, municipalities and sovereigns, as well as structured products and derivatives. Each individual book of business wants to measure its own credit risk, but it is not financially prudent for each to develop or implement its own solution. Furthermore, the group level of the bank would like to evaluate the institution-wide portfolio when calculating capital adequacy so as to account for concentration risk and correlation risk across individual books of business.

The solution is to provide an enterprise-class technology platform that includes a large data warehouse of all obligor information, provides a collection of models as services that are run as composite applications appropriate for each individual portfolio and also allows calculation across all portfolios. This implementation leverages third-party data for the enterprise as a whole, allows the bank to maintain a certain degree of autonomy at the individual business level but, at the same time, meets regulatory requirements to calculate all credit risk across the business in a consistent, reliable and repeatable fashion. The principles illustrated in this example of a large multinational bank apply equally to other types of financial institutions with different portfolio characteristics.

Moody's | K·M·V

Geoff Fite
Chief Operating Officer,
Moody's Analytics
Head of Moody's KMV
T. +1 415 874 6602
E. Geoff.Fite@MKMV.com

Jing Zhang
Head of Research, Moody's KMV
T. +1 415 874 6357
E. Jing.Zhang @MKMV.com

www.MKMV.com