Advances in the modelling of credit risk and corporate bankruptcy: Introduction

Stewart Jones and David A. Hensher

Credit risk and corporate bankruptcy prediction research has been topical now for the better part of four decades, and still continues to attract fervent interest among academics, practitioners and regulators. In recent years, the much-publicized collapse of many large global corporations, including Enron, Worldcom, Global Crossing, Adelphia Communications, Tyco, Vivendi, Royal Ahold, HealthSouth, and, in Australia, HIH, One.Tel, Pasminco and Ansett (just to mention a few), has highlighted the significant economic, social and political costs associated with corporate failure. Just as it seemed these events were beginning to fade in the public memory, disaster struck again in June 2007. The collapse of the 'sub-prime' mortgage market in the United States, and the subsequent turmoil in world equity and bond markets has led to fears of an impending international liquidity and credit crisis, which could affect the fortunes of many financial institutions and corporations for some time to come.

These events have tended to reignite interest in various aspects of corporate distress and credit risk modelling, and more particularly the credit ratings issued by the Big Three ratings agencies (Standard and Poor's, Moody's and Fitches). At the time of the Enron and Worldcom collapses, the roles and responsibilities of auditors were the focus of public attention. However, following the sub-prime collapse, credit-rating agencies have been in the spotlight. At the heart of the sub-prime scandal have been the credit ratings issued for many collateralized debt obligations (CDOs), particularly CDOs having a significant exposure to the sub-prime lending market. In hindsight, many rated CDOs carried much higher credit risk than was implied in their credit rating. As the gatekeepers for debt quality ratings, the 'Big Three' have also been criticized for reacting too slowly to the sub-prime crisis, for failing to downgrade CDOs (and related structured credit products) in a timely manner and for failing to anticipate the rapidly escalating default rates on sub-prime loans. The adequacy of historical default data (and the risk models based on these data) has also been questioned. As it turned out,

historical default rates did not prove to be a reliable indicator of future default rates which surfaced during the sub-prime crisis. Officials of the EU have since announced probes into the role of the ratings agencies in the sub-prime crisis, which are likely to be followed by similar developments in the United States.

Distress forecasts and credit scoring models are being increasingly used for a range of evaluative and predictive purposes, not merely the rating of risky debt instruments and related structured credit products. These purposes include the monitoring of the solvency of financial and other institutions by regulators (such as APRA in Australia), assessment of loan security by lenders and investors, going concern evaluations by auditors, the measurement of portfolio risk, and in the pricing of defaultable bonds, credit derivatives and other securities exposed to credit risk.

This book has avoided taking the well-trodden path of many credit risk works, which have tended to be narrowly focused technical treatises covering specialized areas of the field. Given the strong international interest in credit risk and distress prediction modelling generally, this volume addresses a broad range of innovative topics that are expected to have contemporary interest and practical appeal to a diverse readership, including lenders, investors, analysts, auditors, government and private sector regulators, ratings agencies, financial commentators, academics and postgraduate students. Furthermore, while this volume must (unavoidably) assume some technical background knowledge of the field, every attempt has been made to present the material in a practical, accommodating and informative way. To add practical appeal and to illustrate the basic concepts more lucidly, nearly all chapters provide a detailed empirical illustration of the particular modelling technique or application being explained.

While we have covered several traditional modelling topics in credit risk and bankruptcy research, our goal is not merely to regurgitate existing techniques and methodologies available in the extant literature. We have introduced new techniques and topic areas which we believe could have valuable applications to the field generally, as well as extending the horizons for future research and practice.

The topics covered in the volume include logit and probit modelling (in particular bivariate models); advanced discrete choice or outcome techniques (in particular mixed logit, nested logit and latent class models); survival analysis and duration models; non-parametric techniques (particularly neural networks and recursive partitioning models); structural models and reduced form (intensity) modelling; credit derivative pricing

models; and credit risk modelling issues relating to default recovery rates and loss given default (LGD). While this book is predominantly focused on statistical modelling techniques, we recognize that a weakness of all forms of econometric modelling is that they can rarely (if ever) be applied in situations where there is little or no prior knowledge or data. In such situations, empirical generalizations and statistical inferences may have limited application; hence alternative analytical frameworks may be appropriate and worthwhile. In this context, we present a mathematical and theoretical system known as 'belief functions', which is covered in Chapter 10. Belief functions are built around belief 'mass' and 'plausibility' functions and provide a potentially viable alternative to statistical probability theory in the assessment of credit risk. A further innovation of this volume is that we cover distress modelling for public sector entities, such as local government authorities, which has been a much neglected area of research. A more detailed breakdown of each chapter is provided as follows.

In Chapter 1, Bill Greene provides an analysis of credit card defaults using a bivariate probit model. His sample data is sourced from a major credit card company. Much of the previous literature has relied on relatively simplistic techniques such as multiple discriminant models (MDA) or standard form logit models. However, Greene is careful to emphasize that the differences between MDA, and standard form logit and probit models are not as significant as once believed. Because MDA is no more nor less than a linear probability model, we would not expect the differences between logit, probit and MDA to be that great. While MDA does suffer from some limiting statistical assumptions (particularly multivariate normality and IID), models which rely on normality are often surprisingly robust to violations of this assumption. Greene does stress, however, that the conceptual foundation of MDA is quite naive. For instance, MDA divides the universe of loan applicants into two types, those who *will* default and those who *will not*. The crux of the analysis is that at the time of application, the individual is as if 'preordained to be a defaulter or a nondefaulter'. However, the same individual might be in either group at any time, depending on a host of attendant circumstances and random factors in their own behaviour. Thus, prediction of default is not a problem of classification in the same way as 'determining the sex of prehistoric individuals from a fossilized record'.

Index function based models of discrete choice, such as probit and logit, assume that for any individual, given a set of attributes, there is a definable probability that they will actually default on a loan. This interpretation places all individuals in a single population. The observed outcome (i.e., default/no

default), arises from the characteristics and random behaviour of the individuals. Ex ante, all that can be produced by the model is a *probability*. According to the author, the underlying logic of the credit scoring problem is to ascertain how much an applicant resembles individuals who have defaulted in the past. The problem with this approach is that mere resemblance to past defaulters may give a misleading indication of the individual default probability for an individual who has not already been screened for a loan (or credit card). The model is used to assign a default probability to a random individual who *applies* for a loan, but the only information that exists about default probabilities comes from previous loan *recipients*. The relevant question for Greene's analysis is whether, in the population at large, Prob[D=1|x] equals Prob[D=1|x and C=1] in the subpopulation, where '$C = 1$' denotes having received the loan, or, in our case, 'card recipient'. Since loan recipients have passed a prior screen based, one would assume, on an assessment of default probability, Prob[D=1|x] must exceed Prob[D=1|x, C=1] for the same x. For a given set of attributes, x, individuals in the group with $C = 1$ are, by nature of the prior selection, less likely to default than otherwise similar individuals chosen randomly from a population that is a mixture of individuals who will have $C = 0$ and $C = 1$. Thus, according to Greene, the *unconditional* model will give a downward-biased estimate of the default probability for an individual selected at random from the full population. As the author notes, this describes a form of censoring. To be applicable to the population at large, the estimated default model should condition specifically on cardholder status, which is the rationale for the bivariate probit model used in his analysis.

In Chapters 2 and 3, Stewart Jones and David Hensher move beyond the traditional logit framework to consider 'advanced' logit models, particularly mixed logit, nested logit and latent class models. While an extensive literature on financial distress prediction has emerged over the past few decades, innovative *econometric* modelling techniques have been slow to be taken up in the financial sphere. The relative merits of standard logit, MDA and to a lesser extent probit and tobit models have been discussed in an extensive literature. Jones and Hensher argue that the major limitation of these models is that there has been no recognition of the major developments in discrete choice modelling over the last 20 years which has increasingly relaxed the behaviourally questionable assumptions associated with the IID condition (independently and identically distributed errors) and allowed for observed and unobserved heterogeneity to be formally incorporated into model estimation in various ways.

The authors point out a related problem: most distress studies to date have modelled failure as a simplistic binary classification of failure vs. nonfailure

(the dependent variable can only take on one of two possible states). This has been widely criticized, one reason being that the strict legal concept of bankruptcy may not always reflect the underlying economic reality of corporate financial distress. The two-state model can conflict with underlying theoretical models of financial failure and may limit the generalizability of empirical results to other types of distress that a firm can experience in the real world. Further, the practical risk assessment decisions by lenders and other parties usually cannot be reduced to a simple pay-off space of just failed or nonfailed. However, modelling corporate distress in a multi-state setting can present major conceptual and econometric challenges.

How do 'advanced' form logit models differ from a standard or 'simple' logit model?. There are essentially two major problems with the basic or standard model. First, the IID assumption is very restrictive and induces the 'independence from irrelevant alternatives' (IIA) property in the model. The second issue is that the standard multinomial logit (MNL) model fails to capture firm-specific heterogeneity of any sort not embodied in the firm-specific characteristics and the IID disturbances.

The mixed logit model is an example of a model that can accommodate firm-specific heterogeneity across firms through random parameters. The essence of the approach is to decompose the stochastic error component into two additive (i.e., uncorrelated) parts. One part is correlated over alternative outcomes and is heteroscedastic, and another part is IID over alternative outcomes and firms as shown below:

$$U_{iq} = \beta' x_{iq} + (\eta_{iq} + \varepsilon_{iq})$$

where η_{iq} is a random term, representing the unobserved heterogeneity across firms, with zero mean, whose distribution over firms and alternative outcomes depends in general on underlying parameters and observed data relating to alternative outcome i and firm q; and ε_{iq} is a random term with zero mean that is IID over alternative outcomes and does not depend on underlying parameters or data. Mixed logit models assume a general distribution for η and an IID extreme value type-1 distribution for ε.

The major advantage of the mixed logit model is that it allows for the complete relaxation of the IID and IIA conditions by allowing all unobserved variances and covariances to be different, up to identification. The model is highly flexible in representing sources of firm-specific observed and unobserved heterogeneity through the incorporation of random parameters (whereas MNL and nested logit models only allow for *fixed* parameter estimates). However, a relative weakness of the mixed logit model is the

absence of a single globally efficient set of parameter estimates and the relative complexity of the model in terms of estimation and interpretation.

In Chapter 3, Jones and Hensher present two other advanced-form models, the nested logit model (NL) and the latent class multinomial logit model (LCM). Both of these model forms improve on the standard logit model but have quite different econometric properties from the mixed logit model. In essence, the NL model relaxes the severity of the MNL condition between subsets of alternatives, but preserves the IID condition across alternatives within each nested subset. The popularity of the NL model arises from its close relationship to the MNL model. The authors argue that NL is essentially a set of hierarchical MNL models, linked by a set of conditional relationships. To take an example from Standard and Poor's credit ratings, we might have six alternatives, three of them level A rating outcomes (AAA, AA, A, called the a-set) and three level B rating outcomes (BBB, BB, B, called the b-set). The NL model is structured such that the model predicts the probability of a particular A-rating outcome conditional on an A-rating. It also predicts the probability of a particular B-rating outcome conditional on a B-rating. Then the model predicts the probability of an A or a B outcome (called the c-set). That is, we have two lower level conditional outcomes and an upper level marginal outcome. Since each of the 'partitions' in the NL model are of the MNL form, they each display the IID condition between the alternatives within a partition. However, the variances are different between the partitions.

The main benefits of the NL model are its closed-form solution, which allows parameter estimates to be more easily estimated and interpreted; and a unique global set of asymptotically efficient parameter estimates. A relative weakness of NL is that it is analytical and conceptually closely related to MNL and therefore shares many of the limitations of the basic model. Nested logit only partially corrects for the highly restrictive IID condition and incorporates observed and unobserved heterogeneity *to some extent* only.

According to Jones and Hensher, the underlying theory of the LCM model posits that individual or firm behaviour depends on observable attributes and on latent heterogeneity that varies with factors that are unobserved by the analyst. Latent classes are constructs created from indicator variables (analogous to structural equation modelling) which are then used to construct clusters or segments. Similar to mixed logit, LCM is also free from many limiting statistical assumptions (such as linearity and homogeneity in variances), but avoids some of the analytical complexity of mixed logit. With the LCM model, we can analyse observed and unobserved heterogeneity

through a model of discrete parameter variation. Thus, it is assumed that firms are implicitly sorted into a set of M classes, but which class contains any particular firm, whether known or not to that firm, is unknown to the analyst. The central behavioural model is a multinomial logit model (MNL) for discrete choice among J_q alternatives, by firm q observed in T_q choice situations. The LCM model can also yield some powerful improvements over the standard logit model. The LCM is a semi-parametric specification, which alleviates the requirement to make strong distributional assumptions about firm-specific heterogeneity (required for random parameters) within the mixed logit framework. However, the authors maintain that the mixed logit model, while fully parametric, is so flexible that it provides the analyst with a wide range within which to specify firm-specific, unobserved heterogeneity. This flexibility may reduce some of the limitations surrounding distributional assumptions for random parameters.

In Chapter 4, Marc Leclere discusses the conceptual foundations and derivation of survival or duration models. He notes that the use of survival analysis in the social sciences is fairly recent, but the last ten years has evidenced a steady increase in the use of the method in many areas of research. In particular, survival models have become increasingly popular in financial distress research. The primary benefits provided by survival analysis techniques (relative to more traditional techniques such as logit and MDA) are in the areas of censoring and time-varying covariates. Censoring exists when there is incomplete information on the occurrence of an event because an observation has dropped out of a study or the study ends before the observation experiences the event of interest. Time-varying covariates are covariates that change in value over time. Survival analysis, relative to other statistical methods, employs values of covariates that change over the course of the estimation process. Given that changes in covariates can influence the probability of event occurrence, time-varying covariates are clearly a very attractive feature of survival models.

In terms of the mechanics of estimation, survival models are concerned with examining the length of the time interval ('duration') between transition states. The time interval is defined by an origin state and a destination state and the transition between the states is marked by the occurrence of an event (such as corporate failure) during the observation period. Survival analysis models the probability of a change in a dependent variable Y_t from an origin state j to a destination state k as a result of causal factors. The duration of time between states is called event (failure) time. Event time is represented by a non-negative random variable T that represents the

duration of time until the dependent variable at time t_0 (Y_{t_0}) changes from state j to state k. Alternative survival analysis models assume different probability distributions for T. As Leclere points out, regardless of the probability distribution of T, the probability distribution can be specified as a cumulative distribution function, a survivor function, a probability density function, or a hazard function. Leclere points out that non-parametric estimation techniques are less commonly used than parametric and semi-parametric methods because they do not allow for estimation of the effect of a covariate on the survival function. Because most research examines heterogeneous populations, researchers are usually interested in examining the effect of covariates on the hazard rate. This is accomplished through the use of regression models in which the hazard rate or time to failure is the fundamental dependent variable. The basic issue is to specify a model for the distribution of t given x and this can be accomplished with parametric or semi-parametric models. Parametric models employ distributions such as the exponential and Weibull whereas semi-parametric models make no assumptions about the underlying distribution. Although most applications of survival analysis in economics-based research avoid specifying a distribution and simply employ a semi-parametric model, for purposes of completeness, the author examines parametric and semi-parameteric regression models. To the extent that analysts are interested in the duration of time that precedes the occurrence of an event, survival analysis represents a valuable econometric tool in corporate distress prediction and credit risk analysis.

In Chapter 5, Maurice Peat examines non-parametric techniques, in particular neural networks and recursive partitioning models. Non-parametric techniques also address some of the limiting statistical assumptions of earlier models, particularly MDA. There have been a number of attempts to overcome these econometric problems, either by selecting a parametric method with fewer distributional requirements or by moving to a non-parametric approach. The logistic regression approach (Chapters 2 and 3) and the general hazard function formulation (Chapter 4) are examples of the first approach.

The two main types of non-parametric approach that have been used in the empirical literature are neural networks and recursive partitioning. As the author points out, neural networks is a term that covers many models and learning (estimation) methods. These methods are generally associated with attempts to improve computerized pattern recognition by developing models based on the functioning of the human brain, and attempts to implement learning behaviour in computing systems. Their weights (and other parameters) have no particular meaning in relation to the problems to

which they are applied, hence they can be regarded as pure 'black box' estimators. Estimating and interpreting the values of the weights of a neural network is not the primary modelling exercise, but rather to estimate the underlying probability function or to generate a classification based on the probabilistic output of the network.

Recursive partitioning is a tree-based method to classification and proceeds through the simple mechanism of using one feature to split a set of observations into two subsets. The objective of the spilt is to create subsets that have a greater proportion of members from one of the groups than the original set. This objective is known as reducing the impurity of the set. The process of splitting continues until the subsets created only consist of members of one group or no split gives a better outcome than the last split performed. The features can be used once or multiple times in the tree construction process.

Peat points out that the distinguishing feature of the non-parametric methods is that there is no (or very little) a priori knowledge about the form of the true function which is being estimated. The target function is modelled using an equation containing many free parameters, but in a way which allows the class of functions which the model can represent to be very broad. Both of the methods described by the author are useful additions to the tool set of credit analysts, especially in business continuity analysis, where a priori theory may not provide a clear guide on the functional form of the model or to the role and influence of explanatory variables. Peat concludes that the empirical application of both of methods has demonstrated their potential in a credit analysis context, with the best model from each non-parametric class outperforming a standard MDA model.

In Chapter 6, Andreas Charitou, Neophytos Lambertides and Lenos Trigeorgis examine structural models of default which have now become very popular with many credit rating agencies, banks and other financial institutions around the world. The authors note that structural models use the evolution of a firm's structural variables, such as asset and debt values, to determine the timing of default. In contrast to reduced-form models, where default is modelled as a purely exogenous process, in structural models default is endogenously generated within the model. The authors examine the first structural models introduced by Merton in 1974. The basic idea is that the firm's equity is seen as a European call option with maturity T and strike price D on asset value V. The firm's debt value is the asset value minus the equity value seen as a call option. This method presumes a very simplistic capital structure and implies that default can only occur at the

maturity of the zero-coupon bond. The authors note that a second approach within the structural framework was introduced by Black and Cox (1976). In this approach default occurs when a firm's asset value falls below a certain threshold. Subsequent studies have explored more appropriate default boundary inputs while other studies have relaxed certain assumptions of Merton's model such as stochastic interest rates and early default. The authors discuss and critically review subsequent research on the main structural credit risk models, such as models with stochastic interest rates, exogenous and endogenous default barrier models and models with mean-reverting leverage ratios.

In Chapter 7, Edward Altman explores explanatory and empirical linkages between recovery rates and default rates, an issue which has traditionally been neglected in the credit risk modelling literature. Altman finds evidence from many countries that collateral values and recovery rates on corporate defaults can be volatile and, moreover, that they tend to go down just when the number of defaults goes up in economic downturns. Altman points out that most credit risk models have focused on default risk and assumed static loss assumptions, treating the recovery rate either as a constant parameter or as a stochastic variable independent from the probability of default. The author argues that the traditional focus on default analysis has been partly reversed by the recent increase in the number of studies dedicated to the subject of recovery rate estimation and the relationship between default and recovery rates. The author presents a detailed review of the way credit risk models, developed during the last thirty years, treat the recovery rate and, more specifically, its relationship with the probability of default of an obligor. Altman also reviews the efforts by rating agencies to formally incorporate recovery ratings into their assessment of corporate loan and bond credit risk and the recent efforts by the Basel Committee on Banking Supervision to consider 'downturn LGD' in their suggested requirements under Basel II. Recent empirical evidence concerning these issues is also presented and discussed in the chapter.

In Chapter 8, Stewart Jones and Maurice Peat explore the rapid growth of the credit derivatives market over the past decade. The authors describe a range of credit derivative instruments, including credit default swaps (CDSs), credit linked notes, collateralized debt obligations (CDOs) and synthetic CDOs. Credit derivatives (particularly CDSs) are most commonly used as a vehicle for hedging credit risk exposure, and have facilitated a range of flexible new investment and diversification opportunities for lender and investors. Increasingly, CDS spreads are becoming an important source