Regulatory implications of credit risk modelling

Patricia Jackson a, William Perraudin b,*

a Bank of England, London, UK
b Department of Economics, The Institute for Financial Research, Birkbeck College, Gresse Street, London W1P 2LL, UK

Abstract

This introduction places in context the papers on credit risk modelling contained in the special issue. We explain why credit risk modelling has become such a focus of interest for practitioners and financial supervisors. Even though, as we explain, the current modelling technologies have significant weaknesses, they offer the possibility of major changes in the ways banks are managed and regulated. The main impediment to greater use of these models, especially by regulators, is the difficulty involved in back-testing the risk measures they produce. We suggest some thoughts on how back-testing and other types of model assessment might be performed. © 2000 Elsevier Science B.V. All rights reserved.

JEL classification: G21; G28

Keywords: Credit risk modelling; Bank regulation; Capital requirements

1. Introduction

The papers in this special issue were all presented at a conference on Credit Risk Modelling and the Regulatory Implications organised by the Bank of England.

* Corresponding author. Tel.: +44-171-631-6404; fax: +44-171-631-6416.
E-mail address: wperraudin@econ.bbk.ac.uk (W. Perraudin).

1 Opinions expressed in this introduction are those of the authors and not necessarily those of the Bank of England.
England and the Financial Services Authority in September 1998. 2 The papers are concerned with a new class of models proposed by finance practitioners and academics in recent years. The models in question are designed to assess the risks involved in holding portfolios of credit-sensitive instruments. The emphasis on portfolios (rather than individual exposures) distinguishes these models from the substantial amount of research that had previously been carried out on pricing specific defaultable claims. Thus, for example, correlations between the returns on credit-sensitive instruments are crucial if one’s concern is with modelling portfolio risk, whereas they are irrelevant if the objective is to price a single debt issue.

A survey of several important, current credit risk models is provided by Crouhy et al. (2000, this issue). Broadly speaking, the new models may be classified into (i) mark-to-market, portfolio-theoretic models such as JP Morgan’s CreditMetrics®, or the Merton-style model implemented by the consulting firm KMV, and (ii) default-mode models such as CSFP’s Credit-Risk+® or the approach advocated by McKinsey & Co (PortfolioView®). The default-mode models estimate the distribution of total defaults on exposures in the portfolio over a given horizon. The mark-to-market models estimate the distribution of portfolio value at some future date allowing for credit quality declines, even if they are short of full default, and thereby generate measures of portfolio Value-at-Risk (VaR). An alternative to VaR-style risk measures is proposed by Nakazato (2000, this issue).

One may also classify the models according to the data they employ. Since market prices are generally not available for loan books, making it difficult to construct series of past returns, other kinds of information must be used to proxy changes in value or credit standing. For example, in the CreditMetrics approach, the distribution of the future value of a loan is estimated using the probabilities that the rating of the loan will change and assuming that the yields in the event of such a change are a set of known constants based on forward yield curves. The ratings (which together with spreads proxy for loan value) may either be generated internally by the bank in question or taken from agencies such as Moody’s or Standard and Poors. By contrast, KMV’s approach uses the classic Merton (1974) model to deduce loan value changes from changes in equity market capitalisation and leverage levels. Anderson and Sundaresa (2000, this issue) investigate empirically debt pricing models which generalize the Merton framework.

---

2 The conference was organised in co-operation with the Bank of Japan, Board of Governors of the Federal Reserve System and the Federal Reserve Bank of New York.
1.1. Why credit risk models are important

Credit risk models have attracted attention, first, because they have emerged just as a significant shift has occurred in debt and loan markets. This shift has involved the creation of major new markets in credit derivatives and the unprecedented growth in the pre-existing markets for loan sales and securitisations. For the first time, bank treasurers have been able to manage their credit risks actively, eliminating credit risk “hot spots” in their portfolios and altering their risk exposure as the economic cycle or the bank’s own financial well-being evolve. For banks and other financial institutions wishing to identify and implement risk-mitigating strategies for portfolios of counterparty exposures, the new credit risk models offer extremely valuable conceptual frameworks. The models also facilitate the pricing of portfolios of exposures included in securitisations.

Interest in credit risk models has been further stimulated by the suggestion that such models be used as a basis for banks’ calculations of regulatory capital. The greater marketability of credit exposures combined with increasing focus within banks on the assessment of economic risk has led to strains in the existing regulatory framework instituted by the 1988 Basel Accord. Within that framework, banks must hold capital equal to at least 8% of their private sector exposures. This creates incentives for banks to move exposures, for which their internal capital targets are much less than 8%, out of their books through so-called regulatory arbitrage transactions. The pressures have been particularly great in the US, where by March 1998 outstanding non-mortgage securitisations amounted to the equivalent of around 25% of the risk-weighted loans of some of the major banks (see Jackson et al., 1999). There are also indications of increasing use of the US securitisation market by European banks. The article by Jones in this issue examines the phenomenon of capital arbitrage and provides several examples.

These developments have been a source of concern for three reasons. First, if banks shed their low risk exposures, the average riskiness of what remains will rise, reducing the effective capital buffer implied by the Basel 8% rule. Second, in many cases, regulatory arbitrage transactions reduce capital requirements without a commensurate reduction in the risks the bank faces. For example, securitisations, many of which are driven by regulatory arbitrage, may leave much of the underlying credit risk with the originating bank (see the article by Jones). Third, complex regulatory arbitrage transactions reduce the transparency of banks’ activities, both for regulators and for the market.

The regulatory treatment of activities such as securitisation could be made tougher to discourage regulatory arbitrage but, as the article in this issue by Mingo points out, banks would find alternative means to achieve the same objective through financial innovation. The only effective way to reduce regulatory arbitrage is to align regulatory capital requirements more closely to
banks’ own assessments of economic risks. Different approaches might be taken to achieve this.

The most obvious approach is to divide private and public sector credit exposures into categories (often termed “buckets”) and to impose different percentage capital requirements (or “hair-cuts”) on each category. Such a “bucketing” approach could significantly reduce incentives for regulatory arbitrage if the hair-cuts were correctly chosen. However, it is hard within a bucketing approach to allow for diversification, so banks may still have an incentive to securitise diversified portfolios of loans for which the risk is lower than the regulatory assessment of risk for individual exposures.

The only way to take diversification fully into account is to use the output from portfolio credit risk models as a basis for setting regulatory capital requirements. Hence, the suggestion from many large firms that they be allowed to use internal credit risk models for this purpose. Regulators already have experience of such a system since the 1996 Amendment to the Basel Accord allows sophisticated banks to employ internal market risk models to calculate capital for their trading book exposures.

As we argue below, capital requirements based directly on credit risk models are simply not a practical possibility in the near future. Credit risk models are at a very different stage in evolution from the market risk models which banks are currently permitted to use for calculating trading book capital. Furthermore, credit risk models are far more difficult to back-test than market risk models.

Nevertheless, regulators should explore the degree to which, in the short run, they can make limited use of models. Supervisors are increasingly looking at ways in which they can “fine tune” capital requirements depending on the relative riskiness of a firm’s assets, the strength of its systems and controls and the quality of its management. For example, the UK’s FSA has stated that credit risk models would be reviewed as part of its supervisory process, in order to assess the overall quality of banks’ risk management. Supervisors, just like senior managers, may learn much about a bank’s risk profile by studying the output from a credit risk model. In the longer term, it is possible that credit risk models will be used to set bank capital in a direct fashion.

If regulators do start to place reliance on credit risk models, if only to assess the quality of risk management and therefore the appropriateness of overall capital ratios, then the most urgent requirement is that methods be devised for back-testing models or for otherwise giving banks an incentive to use models in a prudent, conservative fashion. If, in the longer term, banks are permitted to

---

3 This may be difficult since the relative riskiness of different types of credit exposure may vary over time.

4 This was the conclusion reached in the report published recently by the Basel Committee on Banking Supervision “Credit risk modelling: Current practices and applications”, April 1999.
use models to set capital directly, the incentives they face in implementing models will be crucial.

In the remainder of this introductory discussion, we first describe the weaknesses of the current generation of credit risk models, explaining why they are still too primitive in various respects to serve as a basis for direct calculation of regulatory capital. Second, we set out some thoughts on different possible approaches to model recognition, back-testing and penalties. The question of validation and penalties for non-conservative models would become important if the models were to be used as input to supervisory judgments.

2. Weaknesses of credit risk models

2.1. Parameterisation by judgment

In this section, we examine the weaknesses in the current generation of credit risk models. We also describe the current state of knowledge about the performance of these models.

A basic difficulty is that the main models, currently proposed by practitioners, contain parameters which crucially affect the risk measures which are generated, but which must be set on a judgmental basis by the user implementing the model. To take an example, in CreditMetrics, the fraction of idiosyncratic risk assumed for the asset values of obligors is difficult to assess. Typical estimates based on factor models suggest that idiosyncratic risk in equity returns is actually quite high (see Roll, 1988). However, assuming high levels of idiosyncratic risk within CreditMetrics leads to implausible low levels of total portfolio volatility. Hence, analysts working with CreditMetrics generally assume much lower idiosyncratic risk.

To take a second example, within the Merton-style models employed by KMV, a very important parameter is the dividend pay-out rate assumed for firms. According to the model’s assumptions, if firms pay large dividends, the value of their underlying assets drifts towards the bankruptcy trigger and hence their debt is relatively risky. In real life, high dividends often signal that the firm is performing well and are therefore associated with relatively low risk debt. Effectively, the assumptions of the model are too simple to capture this latter effect. So, parameterising the model from historical data on recent dividend payout rates will lead to incorrect results. Hence, the dividend payout rates in the model must be selected in a fairly arbitrary way by the user.

2.2. Risks omitted and data blanks

A second problem with the current generation of credit risk models is that they all have “blind spots” in that there are categories of credit risk which they
simply ignore or correlations between different categories of risk which they take to be zero. For example, in CreditMetrics future spreads for given ratings categories are assumed to be known. The large swings in credit spreads observed through 1998 strongly suggests that this assumption is unjustified. A recent study by Kiesel et al. (1999) shows that, especially for relatively high credit quality exposures, ignoring spread risk within a CreditMetrics-style framework may lead to risk measures which significantly understate the true level of risk.

Current credit risk models also simplify risk measurement by leaving out interest rate or other market risks. This is likely to be a problem if the models are applied to banking book exposures and banking book interest rate risk is not systematically transferred to the trading book (as is a fairly common practice for more sophisticated banks). Jarrow and Turnbull (2000, this issue) discuss these issues and Lotz and Schlögl (2000, this issue), among other results in a theoretical investigation of so-called “market models”, show that default spreads may be related to interest rate levels even if instantaneous default probabilities are constant.

A third problem is what one might term “data blanks”. In many cases, models are applied to types of exposure for which historical data is unavailable. Such models simply assume that the behaviour of such risks is the same as other kinds of exposure for which data can be obtained. For example, data on ratings transitions by US obligors is used to calculate transition matrices which are then employed for non-US obligors. Furthermore, ratings transition matrices based on bond rating data is used in models applied to portfolios of loans. Two of the articles in this issue examine the differences between transition matrices for different countries, industries and stages of the cycle (see the article by Nickell et al. (2000, this issue)) and default rates for bonds versus private placements, e.g., syndicated bank loans (see Altman and Suggitt, 2000, this issue).

2.3. The performance of models

Given the above apparent weaknesses of credit risk models, it is natural to ask what evidence is available about their performance. Up to now, relatively few studies have been completed which systematically analyse credit risk models from an empirical standpoint. The articles by Gordy (2000, this issue) and Crouhy and Mark (2000, this issue) present results from simulations on different models, focussing particularly on the extent to which different models yield similar risk estimates for the same portfolio at one point in time.

Gordy’s study focuses on a comparison of the CSFP model, CreditRisk+, and a simplified version of CreditMetrics in which obligors either default or do not but no other ratings changes are considered. He shows, using simulated data, that various risk measures may be obtained using either and that it is possible to parameterise the models so the levels of these measures are broadly comparable. Crouhy et al. (2000, this issue) compare four different credit risk
models on a benchmark portfolio of 1800 bonds diversified across 13 currencies and covering a wide range of countries, maturities and credit qualities. The VaR estimates they produce are broadly similar, the highest being just 50% larger than the lowest.

The only paper which so far has looked at credit risk models on an out-of-sample basis, comparing risk measures with losses which would have been sustained on actual portfolios, is Nickell et al. (1998). They examine the degree to which two standard credit risk models (CreditMetrics and a Merton-style model like that of KMV) accurately estimate Value at Risk for portfolios of Eurobonds over rolling twelve month periods between 1988 to 1998. Their study is somewhat negative in its conclusions since these models yield far more “exceptions” than they would if they were accurately measuring risk. The models appear to perform poorly in particular for non-US obligors and for banks and financials.

3. Incentive systems for credit risk modelling

3.1. Internal models for market risk

If internal credit risk models are to be used as an important part of the system of banking supervision, it is essential that banks have suitable incentives to employ the models in a prudent manner. To understand the problems involved, it is helpful to consider the use of internal models in the context of market risk, a field in which both regulators and firms have now accumulated a reasonable amount of experience.

Under the alternative approach of the 1996 Basel Accord Amendment, banks may use internal VaR models to calculate capital to cover market risk. From the regulators’ standpoint, a number of important safeguards have been built into the regime. (i) Banks must estimate VaRs according to basic standards. They must use at least one year of data and calculate ten-day VaRs at a 1% confidence level. (ii) Required capital equals the higher of 3 times the 60 day moving average of 1% VaR estimates, or the 1% VaR on the current day. In most cases, the former will be the binding constraint. (iii) Models are back-tested and banks face penalties in the form of higher capital charges if their models underestimate the riskiness of their portfolios.

More specifically, under the back-testing arrangements, penalties are imposed if a bank experiences losses greater than the VaR-losses predicted by the model (such losses are terms “exceptions”) on five or more days in a 250 day

---

5 Many banks calculate ten-day VaRs by simply scaling up VaRs for one-day holding periods by the square root of 10. Hence, one may think of the VaR as being a one-day VaR scaled up by $3 \times \sqrt{10}$. 

period. To illustrate the back-testing regime, if a bank suffers seven losses greater than its 1% VaR measures within a 250 day period, then its capital multiplier would be increased from 3 to 3.65.

This approach to back-testing was criticised by Kupiec (1995) who pointed out that many more than 250 observations would be needed to distinguish accurate from inaccurate VaR models in a statistically reliable way. Nevertheless, the threat of a significant rise in capital requirements is probably sufficient to induce banks to adopt fairly conservative approaches in their estimation of market risk VaRs.

3.2. Testing credit risk models

Back-testing is patently far more difficult for credit risk than for market risk VaR models. One problem is that credit risk models are typically implemented for investment horizons of a year or more. For one-day VaR calculations used in the market risk context, it is relatively easy to construct data sets with hundreds of time series observations. This is clearly impossible for credit risk models. These difficulties are compounded by the fact that confidence levels for credit risk VaRs may be as low as 0.1%. Models which under-estimate risk by a large multiple may appear accurate for many years when the confidence level of the VaR calculation is so low.

A second more subtle difficulty stems from the fact that mark-to-market values are generally not available for the exposures to which credit risk modelling is applied. For example, in the KMV model, the model itself must be used to value loans at different points in time. Hence, out-turns that one might want to compare with VaR-loss-levels are not true losses but merely estimates. Since it is important to test the pricing accuracy of the model, as well as the risk estimates it supplies, such models are difficult to validate.

3.3. Model recognition and subsequent back-testing

How might these difficulties be tackled in a practical programme of model evaluation? One should distinguish between (i) standard models implemented on publicly available data, (ii) standard models employing bank-specific information, such as the bank’s internal ratings, and (iii) internally developed models which use in-house data.

First, consider the analysis that regulators would perform when banks apply for recognition of their models. In case (i) (standard models employing public data), regulatory authorities can run their own versions and evaluate the model’s assumptions and the quality of the data. Evaluations of this kind have already begun. Model performance may be assessed using simulated data (as in the Gordy (2000) study in this issue), or time series information on bond prices (as in Nickell et al., 1998).
It is much harder to perform such analysis for models from categories (ii) and (iii) which employ bank specific data. The difficulties stem partly from the fact that banks’ internal data sets typically cover quite short periods of time and partly from the obvious practical problems involved in implementing numerous different models with various data and basic assumptions.

The article by Lopez and Saidenberg (2000) in this issue suggests that even if data series cover a short period of time, models might be tested by examining the accuracy of their predictions in a cross-sectional sense. While scope clearly exists for such evaluation, the strongly positive cross-sectional correlation between the credit standing of different obligors, and the strong cyclical element in credit risk, suggests that understanding the time series dimension of credit risk is quite crucial.

Second, consider the testing which regulators would carry out after models had been recognised to assess their subsequent performance. Here again, standard models based on publicly available data can be evaluated through a range of simulation-based and time-series-data-based tests. But, again, it would be difficult to perform similar analysis for models which employ bank-specific data and adopt non-standard assumptions.

Even if it were possible to test models satisfactorily, the way in which a particular bank used its models remains important. As in the market risk context, the output of the model would reflect the pre-processing the bank performs before applying the model to the portfolio (for example, netting apparently offsetting risks).

3.4. Cross-bank comparisons

Given the serious difficulties involved in testing heterogeneous bank models, regulatory authorities would necessarily have to find other ways to establish that individual bank credit risk models are reasonably conservative. One possibility is the systematic use of cross-bank comparisons. For example, for models that employ internal ratings, regulators could ensure that different banks rate given individual obligors reasonably consistently. Also, rating transition probabilities could be compared across different models.

Comparisons of this kind would be relatively easy within a given national banking market, but more challenging internationally since the risks associated, for example, with one type of business in different countries might vary significantly. Different industrial structures and insolvency procedures could also lead to different transition probabilities and different loss transitions.

If such comparisons were to be used, it would be important that models have reasonably homogeneous inputs and outputs. Just as in the market risk
context, VaR calculations must be performed according to basic rules (for market risk, 1% VaRs over ten days and using at least a year of data). Regulators could insist, for example, that where internal ratings were an input to the models, rating categories be based on certain principles. The article by Treacy and Carey (2000) in this issue discusses in detail current practice in internal ratings systems employed by large US banks and suggest various features which represent “best practice”.

An important issue is how much standardisation should be imposed on the models – mark-to-market rather than default-mode, a set confidence level, requirements on the amount of data employed, limits on admissible correlations and so on. The more the elements are standardised, the more the models would depart from the firms’ own perception of appropriate risk measurement.

3.5. Bench-marking

A more systematic approach to cross-bank comparisons which might have certain advantages is an approach which one might term “bench-marking”, whereby regulators would require banks to ensure that their models yield specific VaRs and hence given implied capital levels for a series of portfolios which the regulator would propose. As discussed above, most models contain crucial parameters which can only be set on a judgmental basis. A process of bench-marking models would offer a way of setting these parameters at levels which would imply sensible levels of capital for representative portfolios. The regulators could set the capital required for a particular portfolio in relation to best practice in the market (the output of the better models) and their own experience in building and testing models. A view would have to be taken on how conservative the capital buffer should be and therefore the implied default risk for the banks.

An advantage of bench-marking compared with more piecemeal comparisons of components of banks’ credit risk models is that it would impose an absolute rather than a relative standard. To illustrate this, consider what might happen if banks were all required to apply similar ratings to a particular group of obligors. If banks collectively take an over-optimistic view of default probabilities, enforcing comparable ratings would not prevent an erosion in capital.

On the other hand, if bank models were all required to yield capital requirement outputs equal to given levels for particular portfolios proposed by the regulators, then such erosion could not occur. In particular, if the portfolios proposed included one with rating category weights comparable to those of the loan market as a whole, regulators would be able to influence the aggregate credit risk capital of the banking sector. Another important advantage
of bench-marking is that it would avoid the necessity of standardising the parameters of the models, such as the confidence level.

A significant difficulty with bench-marking (which it shares with other forms of cross-bank comparisons) is that it would be hard to apply across several countries. Credit exposures from different national banking markets may be more or less risky and the types of portfolios may differ. The question would then arise as to whether the same portfolios and over-all target capital levels should be employed.

Even within a market, given the different niches occupied by different banks, there might be a question about the appropriate level of capital for a particular bank. A bank specialising in one type of business, say mortgages, might need more capital to back that portfolio than a bank diversified across a range of types of exposure. This would increase the complexity for the supervisors in setting the range of portfolios. Also the VaR would have to be calculated for exposures on which the bank might not have data. This would mean that the supervisors would have to provide a data set.

3.6. Penalty structures

If internal models played a significant role in the regulatory regime for banks, the incentives banks faced to apply them conservatively would be very important. This is particularly true given the difficulties in standardising the components and parameters of the models and in carrying out extensive back-testing.

A somewhat radical approach to capital regulation recently discussed in the context of market risk is the notion of precommitment (see Kupiec and O’Brien, 1997), under which regulated institutions would pay fines if their trading book losses exceeded pre-specified levels but would not be obliged to use any given type of model in their management of risk. Nor would the regulators try to assess the adequacy of the models.

This approach might appear to have advantages in the credit risk context since it permits banks to decide in a decentralised fashion what models are appropriate for their particular circumstances. It is noticeable that banks are already using a wide range of credit risk models, presumably reflecting their different preferences and needs.

The main drawback of precommitment is that the policy is time inconsistent. It is not credible that regulators would impose significant fines on institutions if losses have pushed them close to insolvency. Banks will understand this ex ante and anticipate that fines will not be levied. Since market risk is typically small in relation to total bank assets, banks may make large trading losses but remain financially solid. Banks which make large credit losses are more often
pushed to the brink of solvency, however, and hence the time consistency problems are all the greater.

An additional problem with precommitment is that unless a well-specified credit risk model is employed, it is hard for regulators to observe credit losses in a clear manner. Current accounting approaches to measuring credit losses are simply too opaque to serve as a basis for precommitment-style penalties.

For these various reasons, in any practical credit risk model penalty regime, it would not be possible for the supervisors to refrain from checks on the adequacy of the models and simply rely on penalties, which in extreme circumstances would not work.

Hence, supervisors would have to develop a combination of tests of adequacy and penalties for insufficiently conservative models.

The conservatism of the models and the way in which they were applied to a bank’s portfolio could only be fully assessed by comparing losses made by a bank on a portfolio over say a twelve month period with the VaR estimated for that period by the model.

The supervisors could also check the out-turn on some of the components. For example, whether the transitions seemed consistent with the probabilities assumed in the model at the start of the year. Supervisors would have to take into account the extent to which a bank had changed the profile of its portfolio over the year when assessing out-turns, given that banks could not be expected to calculate the out-turn on a static portfolio.

4. Conclusion

This introduction has set the scene for the articles that follow in this Special Issue on credit risk modelling. The development of new models coincides with and reinforces a potentially revolutionary change in credit markets and banking regulation. As the articles in this issue by Jones (2000) and Mingo...

---

6 A variation of precommitment which might appear more feasible would be one in which instead of levying fines ex post if banks make large credit losses, regulators adjust upwards the bank’s regulatory capital requirement. However, if the cost to equity-holders of (i) a fine or (ii) a requirement that capital be held at greater than profit-maximising levels are equal, either could push the bank into insolvency. So the distinction between fines and higher capital requirements, in extreme circumstances, is not as great as it might appear.

7 In a portfolio-theoretic model such as CreditMetrics, values at the end of the horizon would be estimated rather than directly observed which would potentially create problems. Within default-mode models such as CreditRisk+, the problem is more serious since such models so not attempt to place a market value on the portfolio at different points in time, although such models could presumably be used to generate discount factors. Either way, the terminal value would just be an estimate.
(2000) underline, in the face of these developments, bank behaviour and supervisory practice are evolving rapidly.

We argue above that the weaknesses in credit risk models currently make it inappropriate for them to be employed as the sole determinant of regulatory capital for credit risk. However, supervisors should explore less ambitious roles which credit risk models might play in the supervisory regime. Supervisors, like senior managers, may learn much from the output of credit risk models and they can help to make informed decisions about the quality of risk management and the appropriate capital levels for a particular bank.

A crucial step in this process will be the design of the framework of backtesting, cross-bank comparisons and penalty structures necessary if supervisors are to have confidence that banks are applying their models in a prudent fashion.

References