BAD DEBT PROJECTION MODELS
An Overview of Modeling Approaches

Gerard Scallan

A. Introduction

This paper outlines some ideas that Scoreplus has applied to building financial models of consumer credit portfolios. The primary focus of these models has been to estimate future bad debts. Other applications have included fixing score cutoffs and evaluating the profitability of marketing channels. The models discussed are not standard products. The development and implementation of a model depend on understanding the use that will be made of it, the data sources available and the management priorities of the lender.

Portfolio modeling has got lots of open questions. The questions are both statistical and managerial. Compared with scorecard building, which is relatively standard, this is still an area "under construction." So Scoreplus does not have any pretension to have complete solutions to all the potential issues that can occur.

This paper gives a conceptual overview of the models. It does not cover the mathematical details of the methods, nor give precise project plans.

B. Scope of Models

B.1. Nature of Projections

B.1.1 Projections v. Forecasts

All the models discussed in this paper aim to create projections of future payment patterns on a portfolio, using the payment history observed on the portfolio. They are extrapolations of the payment behavior seen to date. They do not attempt to model explicitly the macro-economic and social changes that clearly affect the level of bad debts (and overall portfolio development). The other principal influence on future levels of bad debt is that of policies and systems under the lender's control - for instance, collections policies. These are not incorporated explicitly in the models.

As far as actual results reflect all these factors, the model outcome is better considered a projection rather than a forecast. For reasons discussed below, making forecasts in this area is both difficult and dangerous. The many sources of uncertainty would lead to forecasts without meaning. The starting point is the projection - which reflects the internal mechanics of the portfolio. Management anticipations (of the economic and market environment) and management decisions (on marketing and account management) modify the basic projection. The result is best qualified as a target or as a budget.

The actual result will be at some distance from this target. The likely range of purely random statistical fluctuation can be estimated from the projection. Beyond this, the environment and policies will certainly change in ways that have not been anticipated. This does not reduce the usefulness of targets. They serve as reference points. Significant departures demand an explanation. They modify the conceptual model used by management to create policy.

The models do allow for sensitivity analysis. In each case, a mechanism allows the model parameters to be adjusted. A judicious choice of parameters can show how the portfolio would react to economic and managerial scenarios. However, management judgment is needed to obtain final results.

B.1.2 Economic Factors

Ultimately, bad debt is largely driven by the economic environment. Macro-economic studies suggest that it is best explained by a combination of interest rates and employment (or unemployment) levels. However, these interact in a complex way with portfolio structure, and are accompanied by important lags. Thus, even if future macro-economic statistics were known with absolute certainty, the portfolio outcome would still be very uncertain. In particular, where a portfolio is marketing-driven, economic projections are inadequate.

The use of economic data for forecasting requires information from a very long period. Bad debts are clearly connected to the economic cycle. Most economists consider that the cycle is irregular, but lasts about seven years on average. However, making conclusions from one economic cycle alone is dangerous. At a minimum, the model should be developed from one cycle and tested on another. This suggests that at least fifteen years' data are needed to develop a useful economic model.

Every cycle is different. In the US, for instance the current economic situation (low unemployment, little inflation, moderate growth and balanced budget) bears very little resemblance to that in the late 1980's or early 1990's. In the credit industry, the expansion in credit card debt has changed the structure of consumer debt significantly. Therefore, a simple transposition from one economic cycle to the next cannot be justified.

Sensitivity analysis is the most practical approach to considering variations in portfolio performance due to economic variation. Macro-economic forecasts are not always accurate and cannot be used without reflection. In practice, management will want to consider how the portfolio might react under several different economic scenarios.

The use of the portfolio projection models discussed in this paper will allow the identification of "underlying" portfolio behavior parameters - i.e. once the effects of the portfolio composition and internal policies have been disentangled, there will remain a residual variation in bad debt from one period to another. The real econometric problem is to explain these "exogenous" variations in macro-economic terms. Such models are likely to be more accurate and more useful to management, than models for the headline portfolio bad debt charge. Therefore, the development of internal portfolio models will go hand in hand with the development of better external economic models.

B.1.3 Management Involvement

Portfolio models do not get management off the hook. Models project out a certain version of history. Judging how new factors will affect portfolio performance must remain a management decision. Models help inform that decision and impose a reality check - are the assumptions made consistent? For instance, if management seeks a 30% reduction in bad debt, what improvement is required in collections roll rates?

Management also plays a key role in model development. The decision as to what is important (and treated by the model) and what is accessory (and ignored by the model) must ultimately remain with management.

A recent review of the use of data analysis in the US credit card industry found a strong competitive advantage to companies where management had a realistic and coherent view of what drives portfolio profitability. On the other hand, it found little difference in the types of models used between leaders and also-rans. The key difference is whether management understands and reacts to model results, not whether the model uses any particular technology. This implies that models must be designed with management use in mind. It also suggests that modelers must devote much care to explaining their models to the key decision makers within the company.

B.2. Data Sources

B.2.1 Portfolio v. Operational Models

Any model is based on data. Poor data leads to poor results. However, the issue is different for operational models (destined to support decisions on individual cases) and portfolio models (where the decision unit is the portfolio or a substantial segment). Operational models will lead to different treatment for particular cases. The actions must be reasonable on those cases. Therefore, it is important that they be based on accurate data.

For portfolio models, the accuracy of the model on each individual case is less important. What matters is that it is right on average and that it is responsive to the decisions it is used to support. Thus, even if data are biased, it can still give useful results.

B.2.2 Account Level Data

The use of account level data allows the integration of portfolio and operational decision making. For instance, if the portfolio is segmented by score, then the effect of a cutoff change can be estimated directly. However, three complications arise from the use of account-level data:

- Quirks in the accounting system (e.g. accrual of interest) often add unneeded complexity;
- Large volumes of data make the model operationally unwieldy; this hinders sensitivity analyses;
- Changes in the details of system operation can invalidate model results.

In summary, account-level data make the model more cumbersome and less robust.

B.2.3 Length of history

Generally, longer histories are needed for financial models than for scoring models. This is because an idea of variations in payment patterns is necessary, to tune sensitivity analysis. Very often, model design is determined by whatever data is available for a three year period.

B.3. Use of Models

B.3.1 Interpretation

Models must not be used blindly. Any model is an approximation to reality, which selects certain key features and eliminates others, less relevant to the questions addressed by the model. For instance, bad debt models consider customer payments, but do not take into account the pattern of dialer usage. The models are applied to systems in a changing environment. They reflect changes in portfolio structure and in mechanics. Changes in the overall environment can invalidate some of them modeling assumptions. Therefore, the user must understand, at least in outline terms, the structure of the model. The model parameters must be updated regularly. Most often, maintaining some involvement from the model developer in the regular operation of the system is desirable.

B.3.2 Model Updates

The goal of portfolio models is to project financial performance. This is dynamic, so the models must be updated regularly. Periodic rebuilds, as used for scorecards, do not give an acceptable match to reality, for most applications of portfolio models.

Tracking mechanisms and model updating mechanisms must be incorporated into the overall model framework. They are not added extras, but an intrinsic part of model management.

B.3.3 Sensitivity

The future is not like the past. Any credit portfolio is subject to "shocks", from internal and external sources. It is neither possible nor desirable for a company's final provisions to be "untouched by human hands". Any model should allow the user to test the reaction of the portfolio to departures of various kinds from the previous system parameters. The value of the model to the organization depends principally on the degree to which it gives insight into how a portfolio works. This can only be done by presenting multiple scenarios.

B.3.4 Support

Portfolio models typically require ongoing management attention and technical support. Top management has a conceptual model of how the business behaves. This model underlies the strategic business decisions. Through an analytic model, the conceptual model is matched against reality. This requires a continuing dialog between technical support and strategic management. Organizations where this works obtain a significant advantage in running their business, in managing a changing environment and making turning their information into profit.

C. Triangular Matrix Projections

C.1. Tracking Triangles

C.1.1 Fundamental Structure

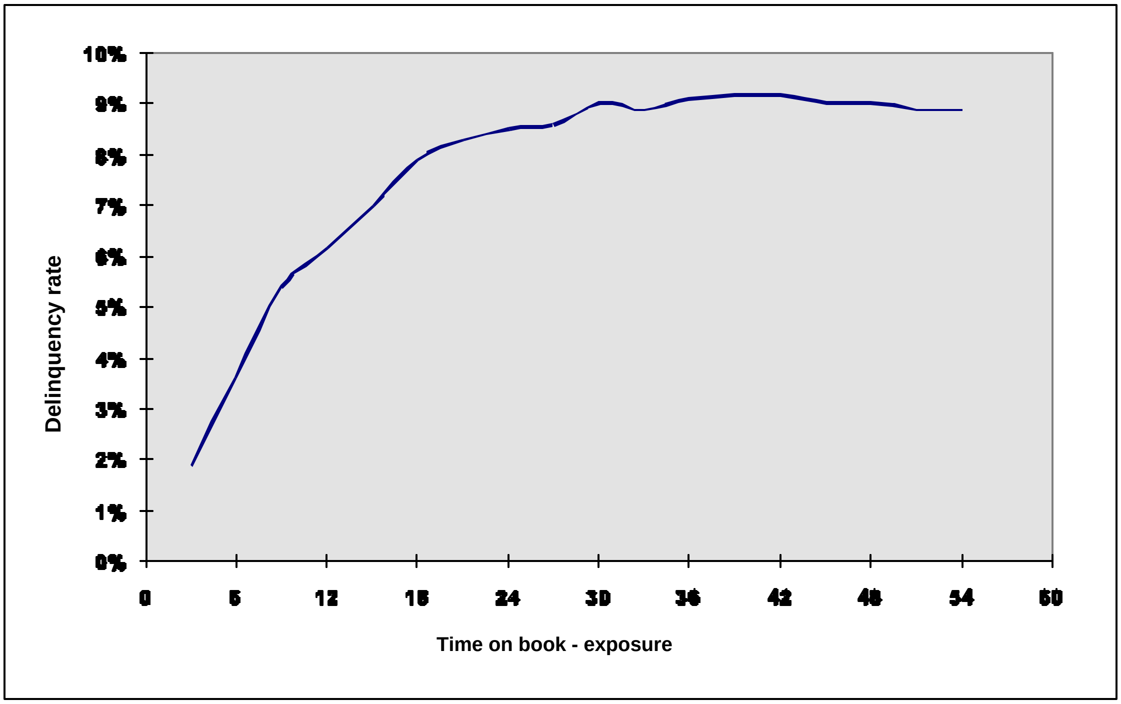
One of the most common portfolio tracking tools is known as a dynamic delinquency matrix. This tracks the cumulative delinquency of accounts opened in a given period, by time on books - see Figure 1. This matrix is also known as a dynamic cohort matrix, cohort performance matrix or delinquency pyramid.

	OPEN DATE						
	93/Q1	93/Q2	93/Q3	93/Q4	94/Q1	94/Q2	94/Q3
Time on Books							
3 months	2.1%	2.4%	1.9%	2.4%	2.2%	2.1%	1.6%
6 months	3.3%	3.7%	3.6%	4.3%	4.3%	4.4%	
9 months	5.0%	4.9%	5.4%	5.0%	6.2%		
12 months	6.5%	6.6%	6.2%	7.2%			
15 months	6.7%	7.6%	7.0%				
18 months	7.2%	8.6%					
21 months	8.0%						
24 months							

Figure 1 - Dynamic Delinquency Matrix

C.1.2 Life Cycle Effect

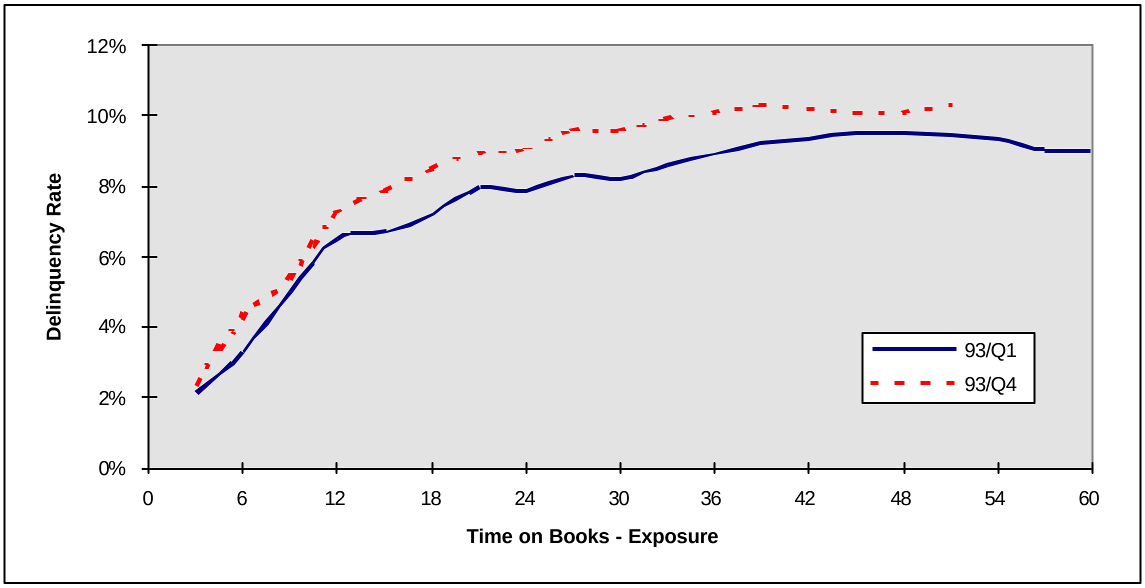
In a dynamic delinquency matrix, the dominant pattern is the "life cycle effect" - the evolution based on time on books. In Figure 1, the maximum delinquency at 3 months is less than the minimum level at 6 months; the maximum level at 6 months is less than the minimum at 9 months etc. Figure 2 illustrates the "life cycle curve". Each curve corresponds to a single column of a dynamic delinquency report.



Here (and most of the time in real life), the life cycle curve is convex - i.e. the increments decrease with increasing time on books, and finish by reaching a plateau or asymptote. The shape of the life cycle curve largely determines the incidence of bad debt. Even if the life cycle were the only phenomenon to take into account, the flow of delinquency would vary, depending on the volumes of business in each cohort.

C.1.3 New Account Effects

In practice, accounts originated in different periods will be of different quality. This is due to variations in marketing or underwriting policy, or to changes in the market. For instance, a drop in score cutoff will normally lead to an increased level of delinquency at each level of exposure. This effect is known as a new account effect. It results in different life cycle curves for business originated in different periods. For instance, in the example given in Figure 1, the delinquency for business written in 93/Q4 is worse, at each level of exposure than the accounts opened in 93/Q1. Figure 3 shows life cycle curves for these two quarters.



The two curves have essentially the same "shape", but are at different levels. Thus the common features (shape) and the differences (position) can be modeled.

In the matrix representation, new account effects are characterized by systematic differences between columns.

C.1.4 Portfolio Effects

New account effects are related to differences in the quality of business written. Sometimes, two cohorts which start the "same" (up to a random fluctuation), will perform in significantly different ways. For instance, a new collections system may shift the performance of the whole portfolio. It will affect different cohorts at different points in their life cycles. This is known as a portfolio effect. Generally, the causes are changes in internal account management (e.g. new collections practices) or in the external credit environment (e.g. changes in bankruptcy laws).

In the matrix representation, portfolio effects are characterized by diagonals significantly different from what might be expected on the base of prior experience. The diagonal corresponds to a fixed observation date.

C.1.5 Model Phenomena

The three effects - life cycle, new account and portfolio - have been illustrated by showing cohort delinquency measures. They can be applied to any aspect of portfolio behavior that is largely determined by a life cycle. Examples include:

- account closures or inactivity;
- conversion to other products (e.g. selling loans to credit card holders);
- time spent in collections (using the date of entry to collections in place of account opening);
- post charge-off recoveries (charge-off date plays the role of account opening);
- limit usage.

The reports and graphs are similar, although the meanings attached to shifts may be quite different. The life cycle approach is a fairly general way of presenting portfolio evolution. To allow for the different applications, the matrices will be called cohort performance matrices from now on.

C.2. Modeling Approach

C.2.1 Fundamental idea

The cohort performance matrix can be modeled by a regression model, with explanatory variables corresponding to the life cycle, new account and portfolio effects. This model is then used to fill the lower part of the triangular matrix, and give estimates of lifetime (or at least, longer term) performance. For each cohort, the model gives an estimate of, for instance, the charge-off level after five years.

C.2.2 Data Requirements

Triangular matrix projections are based on dynamic cohort matrices. Account level data are not used. Therefore, the information needed for modeling is frequently available in the tracking information on the portfolio in question - at least for charge-offs (see below).

C.2.3 Model Construction

The model is built in five steps:

- convert the matrix to incremental increases in delinquency (or other variable of interest); this reduces the autocorrelation between the errors in different cells;
- make a non-linear transformation (frequently something related to the log-odds transformation used in scoring); this reduces the "bunching" toward zero, which occurs in the flat part of the life cycle curve;
- build a non-linear regression model, using terms based on time on books (life cycle effect), on opening date (new account effect), and on observation date (portfolio effect);
- apply the regression model to the cells in the lower half of the cohort performance matrix, to estimate the increments of the regression response variable;
- convert the incremental estimates into cumulative estimates, for each cohort.

Figure 4 shows an example of output from such a model. Observed values are shown in bold. The estimated values are in a light typeface.

C.2.4 Modeling Environment

The volume of data used for a triangular matrix is quite limited. Five years' history, with a cohort for each month, will produce 1770 observations. This number of data-points can be handled in a spreadsheet environment. An efficient method for non-linear regression is needed, to avoid excessive delays. Alternatively, the problem can be readily solved in SAS. The ease of presentation in a spreadsheet moves the balance toward this modeling environment.

C.3. Model Output and Extensions

C.3.1 Presentation of Results

An important advantage of matrix projections is that the output is easily understood. The projection is in the triangular matrix form familiar to many portfolio managers. Predictors used in the model can often be related to shifts in the policies of the lender, in the marketing, underwriting or account management areas.

The matrix diagonals give a projection of future performance period by period. This can be tracked easily. Discrepancies between actual and projected values help identify incipient trends.

The model can also be validated by comparing the actual and estimated values for the "known" cells of the matrix. This provides a convincing non-technical illustration of the model performance (see Figure 5).

C.3.2 Multiple Outcomes

Commonly, accounts vanish from a book in two or more ways. For instance, some accounts go to charge-off, but others pay out normally. A simple application of the triangular matrix for charge-offs risks over-estimating the long term charge-off rate: most accounts pay out before they go to charge-off.

The best way to solve this problem is by building two separate models for the charge-off and pay-out processes. Thus, for each cell in the matrix, the model gives an estimate of the number of accounts that will charge-off and the number paying out. Combining these estimates allows the model to estimate how many accounts are available for either process in the next period.

C.3.3 Financial Values

Ideally, the cohort matrices should be presented in financial terms - e.g. the financial value of charge-offs in a given period, as a percentage of loan advances on the cohort. In practice, the only data available are frequently the numbers of accounts. If so, the account numbers can be converted into financial terms by analyzing the balance at charge-off, or at pay-out, as a percentage of initial advance. However, this introduces a potential source of bias in the estimates - if the distribution of loan terms changes, the distribution of balance at charge-off is likely to change also.

C.4. Model Evaluation

C.4.1 Advantages

The triangular matrix approach is straightforward. Its biggest advantage is its simplicity. It requires very modest amounts of data. The results are easily understood, even for those without mathematical training. The model can be easily manipulated and needs only commonly available analytic tools. Finally, the projections are easily tracked. This means that any feature of portfolio behavior not dealt with by the model can be rapidly and clearly identified.

C.4.2 Limitations

Triangular matrix projections are not a panacea. There are several limitations on the applications suited to this approach:

- The model is most satisfactory for installment lending. For revolving credit there is no natural measure of the "initial advance". This complicates the measures - how should the percentages be expressed. The initial limit may be substituted for loan amount, but the projections are more subject to bias (e.g. from changes in initial limits policy).
- The model reacts only to demonstrated behavior on each cohort. For a charge-off model, the projection will only change when the experience on a cohort is seen to be significantly different from previous cohorts. For instance, a drop in a cutoff would not immediately be reflected in the bad debt projection for the new cohort. Thus, the model does not use all the available information.
- Uncertainty about the ultimate outcome is greatest on recent cohorts. The projection is based on very few observations for these cohorts. Therefore, they have little influence on the regression coefficients. However, this is where the exposure - and likely future losses - is greatest. This problem exists with other modeling approaches, but is particularly acute with the triangular matrix approach.
- Extending the technique to customer modeling is difficult. The triangular matrix is most effective in modeling phenomena which depend on time on books. For a complex customer relationship, this may be difficult to identify. At any given time on books, a greater variety of situations must be taken into account.

C.4.3 Practical Experience

This method has been used since 1993 to track post write-off recoveries with satisfactory results. The projections have performed well and have given considerable insight into portfolio performance. However, the "product" is particularly simple. The projection for cohorts has identified trends reliably from a very early stage, and has obtained wide user acceptance.

The models have recently been restructured and extended to estimating charge-offs on active portfolios. This has considerably improved the modeling environment. Initial results are plausible and have worked well over short periods.

D. Markov Model Projections

D.1. Markov Matrices

D.1.1 Roll Rates

The roll-rate approach has been used for many years to formulate bad debt provisions. Roll-rate models calculate the probability for accounts at each delinquency level of "rolling on" to the next level. For instance, 20% of accounts one cycle delinquent may become two cycles delinquent the following month. Accounts pass though a fixed succession of states before charge-off. By multiplying the successive roll rates, the proportion of accounts that will go to charge-off can be estimated.

The roll-rate approach works well for simple products and simple portfolio management strategies. It breaks down when accounts can "skip" levels, or move backwards. The Markov Model is a generalization of roll rates.

D.1.2 Markov States and Transitions

Instead of looking at delinquency levels, the Markov Model considers "states". For instance, a state could be defined as "Current Delinquency = 0, Behavior Score 500-519, Balance/Advance ≤ 50%" or "3 cycles delinquent, Behavior score 400-439". The state definition must be sufficiently detailed to capture the essentials of the likely future behavior of the account.

Accounts move from one state to another in each period - normally a month or quarter. This is called a transition. Not all combinations are possible - for instance, it may not be possible to move from a cycle delinquency to cycle 4 delinquency. A special state corresponds to closed accounts and another to charge-off cases. Figure 5 shows a set of state definitions for an installment loan portfolio.

D.1.3 Markov Matrices

For the accounts in a one state at the start of a period, there is a certain probability of being in any given state at the end of the period. These probabilities are summarized by a matrix, called a transition matrix. By multiplying out the transition matrix, the movement in the portfolio over any number of periods can be simulated. This multiplication is the basic mechanism by which the Markov Model creates portfolio projections. Figure 6 shows a very simple example of movement through one period.

D.1.4 Markov Assumption

A Markov model simulates the overall development of the portfolio through time. The Markov state must contain enough information to determine the likelihood of moving to any other state in the next period. In other words, all the accounts in a given state must have the same probability of moving to another state in the following period. Therefore, the state definition must be sufficiently detailed to encapsulate the essential "drivers" of account performance. Many companies have applied the methodology with an insufficient set of states. This leads to unstable results over the long term.

D.1.5 Financial Structure

In portfolio projections, financial estimates are critical. The basic Markov matrices model the movement of numbers of accounts. These are complemented by a set of matrices which translate the account numbers into financial volumes. The structure of these matrices depends on the product structure. Typically, they will account for interest charged, new spend (if appropriate) and payments received on accounts moving from one state to another. For instance, payments may account for 3% of outstanding balances on accounts moving from state 13 to state 17 during the period.

These complementary matrices give a set of matrix equations that propagate the portfolio summary from one period to the next. The portfolio summary gives a breakdown of the portfolio by state. Most often, the summary includes for each state:

- number of accounts;
- principal balances;
- interest balances;
- spend in period;
- interest and charges in period;
- payments in period.

D.2. Modeling Approach

D.2.1 Definition of States

In building a Markov Model, the first step is to define a satisfactory set of states. This must meet statistical criteria - ensuring that the states allow the Markov assumption to be met. At least as importantly, the states must satisfy management criteria. Key categories required for management reporting - e.g. early collections - must correspond to groups of states. The definition is therefore a complicated task.

One of the main benefits of the Markov approach over the older roll rate approach is the ability to distinguish levels of risk among the up-to-date accounts. This part of the portfolio must be divided into appropriate states based on probability of delinquency and of early repayment. This feature helps identify changes in underlying portfolio risk much more rapidly.

The methodology for state definition uses decision tree techniques. These must be manipulated with care to take account of management structures and of potential instability due to policy changes.

D.2.2 Role of scores

A behavior score is not necessary to build a Markov model. However, a score is useful in identifying an adequate and economical set of states. The score summarizes the information available about the medium-term risk of an account. Therefore, it is a powerful indicator of what the account behavior will be in the next month or quarter.

Behavior scores are not sufficient to define the Markov states for three reasons:

- Most often, the scores are calculated only on a subset of the accounts - for instance, the severe delinquents are not normally scored. By definition, a Markov model covers a complete portfolio.
- The score summarizes the risk information in the portfolio. A Markov model must also take into account spending and repayment patterns - e.g. probability of early repayment for a loan.
- The categories needed for meaningful management reporting may not correspond to score bands. Even with the same level of risk, up to date accounts must be distinguished from delinquent accounts.

D.2.3 Data Requirements

The matrices which drive the model are calculated from observations of portfolio history. They should be based on observations from an extended period - at least twelve months, by preference twenty four months. These matrices summarize the behavior patterns of the portfolio. The model itself only requires the information at the "matrix" level - i.e. the set of accounts in one state at the start of the period, and in another at the end of the period. Once states have been determined and the accounts summarized into state level information, retaining the account level information is not necessary.

D.2.4 Estimation and Testing

The model is estimated by averaging the observed matrices over several successive periods. The adequacy of the model is tested by:

- taking the state of the portfolio at a point in time;
- moving the portfolio from one period to the next by applying the matrices - both to account numbers and to balances;
- comparing the estimated portfolio breakdown at the outcome date to the actual portfolio results.

If the estimated and actual numbers do not match adequately, then two modifications are possible:

- modify the state definitions, to reflect the Markov property better;
- change the structure of the financial matrices to take account of the cause of errors.

A principal difficulty with the Markov Model method is the lack of algorithms for identifying the appropriate changes. The modeler must use his or her knowledge of the portfolio and of the model mechanics to propose changes. These modifications are then tested.

D.2.5 Updating

The matrices driving the model are updated as new data becomes available. Thus, the model always reflects the most recent behavior patterns on the portfolio. The model is tracked by comparing the previous forecasts with actual output. This should lead to a validation of the state definitions from time to time.

D.2.6 Modeling Environment

The steps to create the matrices are:

- Match two account files, from the start and end of the period;
- For each account, calculate the state at the beginning and end of the period;
- For each combination of start and end-state (corresponding to a cell of the matrix), cumulate the values required for the matrices - number of accounts, opening balances, payments etc.
- Transform the values to match the definitions used in matrices - e.g. average opening balance for each account in a matrix cell.

These steps can be done in SAS, or in any other computer language. Once the matrices are available, the matrix computations are best carried out in SAS IML, Mathematica or another modeling language which uses matrix arithmetic. For simple models, the matrix computation can be done in a spreadsheet such as Excel.

D.2.7 New Accounts

A credit portfolio is constantly renewed. A substantial part of next year's outstandings comes from accounts not yet opened. So, any portfolio projection must take account of the volume and profile of future account openings. This is done by injecting new accounts and balances. These may be isolated in certain states. If the projection is to reflect the changes in quality of new business, there must be several states into which new accounts can enter, with different delinquency probabilities.

D.2.8 Scenarios and Libraries

Over time a library of behavior patterns accumulates. These can be combined to create scenarios. For instance, if the behavior patterns of early 1996 are applied to today's book, the future development of the portfolio can be estimated assuming a different economic environment.

It is also possible to introduce "artificial" modifications to behavior patterns. For instance, a new collections system might be expected to improve roll rates on early collections. The Markov model can be used to estimate the cash flow, profit and bad debt consequences of such an improvement. In this way, the Markov model can help evaluate the costs and benefits for new policies and systems.

D.3. Model Output and Applications

D.3.1 Presentation of Results

The simplest way to use the Markov model is to project forward the current state of the portfolio, with or without the introduction of new business. This results in a set of "state vectors" for each future time period, giving balances, account numbers, payments received etc. From these, the key statistics for the portfolio projection can be computed - e.g. total outstandings, interest income, collections balances, charge-offs. Discount rates can be applied if appropriate.

Figure 6 shows the summary of output from a typical simple Markov model for installment loans. It traces the future portfolio development over a three year period. Typically, results are not presented for the individual states, but at a higher level of aggregation - all up to date accounts are grouped together, for instance. Management decides the level of aggregation and which numbers to report.

Figure 7 shows the analysis of three alternative scenarios for portfolio development, over a three-year horizon. This analysis allows management to identify sensitivity to external environment or to internal policies in financial terms.

D.3.2 Provisioning

Credit companies are required to provision future bad debts arising from current exposure. This can be done readily with a Markov model. For accounts in any given state, the long term outcome can be estimated. This is done by running a projection only on the accounts currently in that state. The projection will calculate the future charge-offs arising from these accounts. A variety of financial assumptions can be used - e.g. suspend interest on certain states, disregard accounts which return to order before ultimate charge-off.

D.3.3 Treasury Planning

Markov models project the cash received for each state and for each time period. They also give figures for balances outstanding. These can be used to identify the treasury requirements for each period in the future. Sensitivity to various parameters can be tested by varying the matrix coefficients - e.g. early repayment. Where the portfolio does not behave as expected, the departures can be quickly identified and the consequences estimated.

D.4. Model Evaluation

D.4.1 Advantages

Markov models provide a coherent framework for addressing broad questions of portfolio models. They integrate financial and account number models to project the overall portfolio balance sheet and cash flow. The models help translate improvement or deterioration of the portfolio into financial terms. This builds a bridge between the risk management and financial measurements of portfolio performance.

The models can readily be formulated to deal with a variety of products. In particular, they deal well with revolving and checking accounts, handling both credit and debit balances (in different states).

The use of a spreadsheet interface makes it easy to play with different scenarios and to present results clearly. This increases their influence on portfolio management. However, use has remained restricted to "experts". There has been little direct use of models by users who have not been involved in model development.

D.4.2 Limitations

State definitions are at the root of most of the limitations on Markov models:

- Defining "good" states is an art. It draws on statistical tools (CHAID, CART), but also requires experience and judgment. This creates a bottleneck in system development: the solution to lack of fit in model results is to redefine states.
- Groups which do not correspond to fixed states - e.g. all loans over $15,000 - are difficult to handle within the model. They can be mapped to states by a "correspondence matrix", but there can be anomalies in the results.
- There is no fixed endpoint for a portfolio. The model assumes that a certain proportion of accounts in each state close in each period. This continues indefinitely. The model does not take account of the scheduled end-date for installment loans. Since most loans do not go to full term, this problem has been more theoretical than practical.
- The reconciliation of accounting figures from one period to the next is painstaking. For example, a detailed knowledge of the accounting system is needed to distinguish adjustments from payments. This potentially makes the model vulnerable to modifications in accounting conventions used.
- The financial parameters (interest rates, charges) are derived from observed data, not entered as system parameters. This makes them difficult to vary. The overall level can be varied, but precise control is not possible.
- It is difficult to get users to use the model to its full extent. This depends on understanding the model, seeing how it is relevant to their problem and having ready access to the model tools.

D.4.3 Practical Experience

Markov models have given robust results on current account and installment loan portfolios over at least five years, even in quite extreme economic environments, such as the UK recession and boom in the early and mid-1990s. Short term behavior patterns (say one quarter) give quite volatile results, but are very effective in identifying the reasons for any change in portfolio performance.

Once a satisfactory set of state definitions has been found, the results appear stable and reliable over a long period. On the other hand, the complexity of the state definitions has been a barrier to user understanding and hence to broadening the use of the model.

Like any other tool, a Markov model must have a strong advocate within the business. Senior management must remain involved in the use of the model if it is to realize its potential impact on the business. The provision of an Excel interface to the model has not been sufficient to overcome the resistance of users.

E. Portfolio Simulation Models

E.1. Simulation Approach

E.1.1 Motivation

The Markov model provides an attractive framework for bad debt projection and for portfolio management. In particular, it allows many risk management issues to be translated into financial terms. However, some important policy questions cannot be formulated in such a model.

For instance, questions relating to new business are poorly addressed by a Markov model - e.g. what is the most profitable score cutoff, or should certain marketing channels be developed? The Markov model does not give enough detail in the relevant areas of the business. Pricing and project design are other areas where the Markov approach does not give satisfactory answers. Finally, organizations with complex customer relationships find it difficult to summarize them into a succinct state definition.

Customer-level simulation is the obvious response to these difficulties. It permits a modular approach to customer behavior. From a very basic starting point, further detail can be added as the policy issues addressed demand.

E.1.2 Lender and Customer Models

Customer behavior in most credit products is very simple, and can be reduced to a few dimensions. For instance, on installment credit, customers make or fail to make a monthly payment. The customer may also decide to pay off the total loan, and in this case start another loan. Finally, a customer can declare bankruptcy. Thus, customer behavior in a given month can be considered as a low-dimensional vector. The behavior is - from the lender's point of view - random, but obeys certain probability rules.

The lender responds to the customer's behavior by taking collections action in case of non-payment, or may offer a top-up loan. These actions are deterministic - there is a finite set of rules which determine the lender's behavior.

A simulation model for this situation is a computer program which emulates the customer's actions (using random numbers and probability distributions), and the lender's response (using fixed rules).

E.1.3 Other products

Credit cards are somewhat more complex than installment loans. The customer spends on merchandise and takes cash advances. He or she also makes payments (or not). As a result, the customer balance goes up or down and the lender takes appropriate actions (limit increase, collections letter). There are at least three dimensions to customer action, and these actions (e.g. merchandise spend) are in a continuous space, not a discrete (pay/no pay) space. Nonetheless, these actions can also be modeled using regression.

For a check account, the issues are similar - the customer communicates by spending money (cash withdrawal, issuing checks, making transfers, paying direct debits) or depositing money. Internal traffic with the customer's savings or revolving loan accounts may also occur. These can also be modeled by scoring and regression models.

E.1.4 Local Scores

Simulation looks at the individual events which constitute customer behavior. The simplest unit is a single month's activity. For each dimension of customer activity, a "sufficient statistic" is computed. This summarizes all the information from the customer's history to estimate the probability distribution of the customer's actions. In the case of discrete actions (e.g. pay / no pay, bankruptcy or early repayment), the probability is estimated through a score. This is called a local score, because it concerns an event in the very near future - will the customer pay or not this month, will the customer close the account this month. It differs from "classic" scores which look at the probability of "bad" behavior over at least a six-month horizon. The local scores are driven by the account opening score and the behavior score, among other predictive elements - in fact everything that is relevant to the probability of the action in question.

E.1.5 Scores and Random Numbers

The simulation works by taking one month at a time. A local score is calculated for each dimension of the customer's behavior. A random number is generated, between 0 and 1. If the number is less than this probability, then the customer is considered to have made a payment, otherwise the customer is considered to have missed the payment. Accounting entries are generated as a function of the simulated action.

For continuous actions (e.g. level of spend), a slightly different mechanism is used, with normal random numbers. However, the basic principle is the same.

E.1.6 Future Histories

Simulation generates the basic events which drive the account history. These events generate records in the same format as the actual account histories. This is called a "future history". It represents a randomly chosen version of what the customer may actually do. Any reports which are produced from the "real" account files can then be emulated on the simulated future histories.

Provided the simulated accounts are sufficiently numerous, the future history will be a good approximation to what will actually happen (given the assumptions which drive the simulation model of individual behavior). The statistical theory behind this is the "Law of Large Numbers".

E.1.7 Garbage In, Garbage Out

Simulation is not a magic wand. It is a mechanism for integrating micro-models (of the most basic elements of customer behavior) into a macro-model (of overall portfolio behavior). If the micro-models are not a good representation of customer behavior, then the macro-model will not be a good model of the overall portfolio. This can be resumed as garbage in, garbage out.

In practice, this pitfall can be turned to advantage. Frequently, "garbage out" is easier to recognize than "garbage in". Incoherent results at the portfolio level indicate faults in the models of individual elements or in the model of their interactions. This helps "debug" the local scores.

E.2. Model Development

E.2.1 Model Data

Simulation models are data hungry. The construction of local scores (or regression estimates for continuous variables) requires detailed histories of prior account performance. Moreover, all elements needed for the estimation of local scores in future months must be generated at each monthly step. Designing a practical simulation scheme requires experience and experiments.

Local scores are built from a database of account histories. The most important elements in local scorecards are "classic" scores - new business scores, behavior scores, bankruptcy scores. The presence of effective scorecards makes it much easier to build a reliable simulation model. Other elements are needed to model the non-risk aspects of behavior - account spend, early closure etc.

The first step in constructing a simulation model is to create a database of account histories, similar to that used for behavior score construction. From this database, observations for local scores can be extracted.

E.2.2 Modeling Environment

A flexible modeling environment is essential for simulation models. Multipurpose score and regression models are needed. SAS or a similar environment is most appropriate for the construction and for the regular use of a simulation model. To date, simulation models are not sufficiently well established to make a good user interface possible.

The outcome from simulation models can best be presented as future histories, and analyzed with the tools (data management and reports) used for portfolio tracking. This creates a large volume of data, particularly where different scenarios are compared. In addition, large samples (up to 100,000 simulated cases) are required to converge to stable outcomes. Simulation therefore demands considerable computing resources - notably computer storage.

E.2.3 Local Score Models

Local scores are the backbone of simulation. Changes in customer behavior patterns are taken into account by modifying the local score. For example, a deterioration in the economic environment will change the relationship between behavior score and the probability of missing a payment. This should be reflected in the local score model. For this reason, local scores need to be re-estimated regularly, on the basis of recent portfolio behavior. This should be done at least once every six months.

E.2.4 Scenarios

Portfolio behavior varies over time. Different versions of local scores, from different time periods, allow scenarios of portfolio behavior to be worked out, reflecting variations in economic and market conditions. It is also possible to create *ad hoc* scenarios - what happens if the incidence of early payout increases by 10%. The problem arises in structuring the various scenarios. The potential variation is so rich that it is difficult to parameterize the possible scenarios. This can lead to problems in administering the simulation system.

E.2.5 Modular Modeling

Not all aspects of customer behavior are of equal interest. For instance, the development of a model may focus initially on bad debt projections. Operational costs are not included. It is possible to add a module of operational cost - customer service, account maintenance, documentation - at a later date, without modifying the existing model.

Modularity has its limits. Suppose the primary focus of the model is bad debt on installment loans. It is nonetheless important to include a rough model of early payout, as this will affect the debt outstanding and subject to bad debt. The early payout model need not consider all determinants of early payout. It must, however, give a realistic model of the overall incidence of early payout, and relate it to account risk measures (scores).

E.2.6 New Scorecards

New accounts must be generated as part of the overall portfolio simulation. These do not have a prior behavioral history. The principal information available about them is their new business score and financial details - loan amount or limit, loan term etc. These elements can be taken into account in the construction of local scores.

When new scorecards are introduced, their relation with local scores cannot be estimated from data. This problem can be tackled in two ways:

- reconstructing local scores from the scorecard development sample;
- establishing an equivalence between new and old scores, through a standard score-odds relationship.

This simulation of the financial impact of new scorecards helps design more effective strategies for the use of new scorecards.

E.2.7 Customer-level models

Simulation models can be applied to customer-level as well as to account-level behavior. Where there are significant interactions between accounts, local scores can use the history of other accounts to model customer actions. This permits the model to address issues such as loan top-ups, or relationship between checking accounts, loans and credit cards.

E.2.8 Model Testing

Simulation models should satisfactorily reconstitute overall portfolio performance. There are three critical measures:

- overall level of portfolio outstandings;
- level of charge-off balances;
- score-odds relationship over the horizon of classic scoring systems.

The importance of the first two measures is obvious. The third measure (score-odds) reflects the importance of score as a measure of risk. If the simulation model cannot satisfactorily reproduce the behavior of the scoring systems (whether new business or behavior scores), then it has not properly captured the mechanics of portfolio movement.

E.3. Output and Applications

E.3.1 Future Histories

The immediate output from a portfolio simulation is a set of "future histories". These give a version of future portfolio performance consistent with the "micro-model" of customer behavior underlying the simulation. These future histories can be analyzed in any number of ways.

Simulation output can be presented in the same format as Markov model output (see Figure 6). However, the simulation approach models the behavior of individual accounts and customers. Thus, accounts can be grouped in any way that seems appropriate.

E.3.2 Charge-off Projections

Once future histories are available, charge-off projections can be produced. This is done by cross-tabulating various categories of account against transfers to charge-off during the outcome period of the simulation. This gives greater flexibility in provisioning and in portfolio budgets than any other method. Groups of interest include:

- Score bands (new business or behavior scores);
- Marketing source codes (for analyzing profitability by source);
- Region or branch;
- Current delinquency level;
- Type of agreement (loan amount, term ...).

E.3.3 Cash Flows and Pricing

Simulation scenarios can vary the financial treatment of accounts. For instance, the effect of delinquency charges or interest rate changes on cash flow and on losses can be assessed. However, changes in pricing would probably change customer behavior. The scenario should allow for these changes. Once again, the principle remains that simulation integrates micro-models (based on assumptions or actual evidence) to give an overall portfolio picture.

E.3.4 Treasury

As with Markov models, simulation can be used to estimate funding requirements for a portfolio. Treasury models arising from simulation can be tested and varied more easily than those from Markov models.

E.3.5 Score Cutoffs

The goal of score cutoffs is to accept profitable business and turn down unprofitable business.

Using simulation models, the portfolio manager can model the sources of profit on loans or credit cards. The sensitivity of profit to assumptions (e.g. recovery rates on charge offs, or the effectiveness of collections action) can be estimated.

E.3.6 Behavior Scoring Experiments

The delay and difficulty of analyzing experiments have hindered the effective use of behavior scoring management systems. Simulation models provide a mechanism for translating short-term results (e.g. limit utilization) into financial terms and extending them over a longer horizon. Simulation is not a miracle cure, but it should contribute to more effective use of behavior scores. In conjunction with well-designed experiments, simulation can give valuable information about customers' reaction to different policies.

E.4. Model Evaluation

E.4.1 Advantages

Portfolio simulation models were developed in response to limitations on Markov models. They provide an overall approach to portfolio modeling. Simulation allows different models to be integrated. Anything which can be done with a Markov model can be done with a simulation model, and with greater power and flexibility.

Simulation has given promising results in setting cutoffs and in modeling overall portfolio cash flow. There are many other applications which remain to be developed.

Simulation models can integrate the results of all kinds of more specialized models to give an overall portfolio picture. They help translate the results of low-level "micro-models" into overall financial terms - in other words, into the language of strategic portfolio management. For this reason, they are likely to become a central part of credit portfolio management in the next decade.

The ability to vary financial parameters and study the consequences is one of the principal advantages of simulation models. The wider use of risk-based and competitive pricing will create a demand for models which optimize policies in more complex environment than today's.

E.4.2 Limitations

Simulation is a new approach to portfolio modeling. It is not yet mature, and there are few practical tools for model development and management. It is still a general approach rather than a specific methodology.

This generality means that the development of a simulation model is open-ended. This makes project management difficult. Certain aspects of customer behavior may emerge as critical during model development. They must then be included in the model.

These drawbacks are most acute for complicated products, with substantial customer interaction. For instance, models of credit card spending are rudimentary compared to risk scorecards. There is a danger that overall portfolio models will be as weak as their weakest link.

E.4.3 Practical Experience

Simulation models have been applied to installment loan business with some success. The customer's behavior is low dimensional and it is relatively easy to measure the fit between simulated and actual results. However, even in this area, the models have not been run for long enough to identify all potential snags.

The lack of a "standard" user interface means that simulation models must be treated as a research area rather than as a standard decision-support tool. Each use of the model requires support from the model builder.

Simulation models are certainly tools for the future. Over the course of the next ten years, they are likely to facilitate the integration of models into the policy making process of credit grantors. However, they represent a promising research area rather than a proven technology in the short term.

F. Comparison and Conclusion

F.1. Comparison

F.1.1 Data Requirements

There is an essential difference between the Triangular Matrix approach on the one hand, and the Markov and simulation models on the other. Triangular models have a predetermined structure. Everything is modeled in terms of opening cohorts. They require very modest amounts of data, in a fixed format. On the other hand, the Markov model and the more general simulation approach, are much more flexible in their model of portfolio structure. The price for this flexibility is the need to access account-level data, and the use of scores.

F.1.2 Accuracy of Projections

A proper test of predictive performance would require parallel models on a large number of portfolios through a variety of economic and market environments. So no definitive answer can be given to the question. However, some general comments are valid.

In a stable situation, there is likely to be little difference between the models. Indeed, if the portfolio is of constant size and quality, the best estimate will be that one year's bad debt performance will be much like another's. Therefore, the question is how a model reacts to changes in portfolio structure and behavior.

All three models will take into account changes in portfolio composition, notably growth in outstandings. When a portfolio grows, new accounts are larger proportion of total balances. New accounts generate little bad debt immediately, but over a one to two years have higher levels of bad debt than more seasoned accounts. So bad debt declines at first, then increases (as a percentage of outstandings). This phenomenon is well reflected in all three models.

A change in the quality of business is more rapidly reflected in Markov models than in Triangular matrices. Triangular matrices do not take into account variations in the performance of up to date and mildly delinquent accounts. In the first months, the model is dominated by the behavior prior to the change. Hence, their reaction is somewhat slower than the other models. Markov models tend, if anything, to overreact to short term changes in portfolio dynamics. This helps concentrate attention on the modifications in portfolio behavior, but needs to be damped to produce plausible forecasts of future bad debt flows. The same problem can affect simulation models, if the sample period used to construct local scores is too short.

F.1.3 Economic Factors

None of the models explicitly takes external economic factors into account. On the other hand, in all models, it is possible to include a "feel good" or "feel bad" factor, which modifies the future behavior patterns in one direction or another. The performance of the portfolio can be calibrated against the observed variations in the past, to determine reasonable values of these factors.

For each model, the modifications apply to the basic units of the model - so for the Triangular Matrix approach, the factors apply to opening cohorts, to particular ages of account or to observation dates (columns, rows or diagonals). For Markov models, the factors can change any subset of states. Thus, if care is taken in defining states in the first place, a large degree of flexibility can be obtained. For instance, in the 1980s, overall portfolio performance was often strong, but particular groups were subject to bankruptcy and post charge-off recoveries suffered from high levels of unemployment. Simulation models are by far the most flexible. Since each individual account is treated individually, the model can differentiate in any way desired. The problem is to identify what modifications are justified.

F.1.4 Applications

The Triangular Matrix model is designed to produce regular projections of bad debt. In broad terms, it also identifies the presence of changes in portfolio quality, and reflects the quality of new business from any given time period. However, it does not allow the causes of any changes to be identified.

The Markov model is much more analytic, and gives a more complete picture of portfolio evolution. The model can produce an overall projection of portfolio evolution, including cash flow, interest income and bad debt, and indicate how this would vary with differing assumptions. This makes it a much more powerful tool for high-level portfolio management.

The simulation model goes further than the Markov model. Everything that the Markov model can do can be done equally well by the simulation model. In addition, the model provides a framework for detailed modeling of policy changes. Thus, the simulation approach provides a bridge which ensures consistency between detailed policy development (e.g. changes in collections letters) and the strategic direction of the portfolio.

F.2. Conclusion

F.2.1 Management Role

The primary purpose of a bad debt model is to facilitate more effective management of a credit portfolio. So, the best system is that which will best support management. The choice depends on the management mechanisms and culture of the organization.

The results from any model must be looked at carefully. They are not infallible predictions of how the portfolio will perform, merely extrapolations of existing behavior. Judgment is necessary to take account of changes which may invalidate past experience as a guide to future performance.

F.2.2 Future Development

Credit management is becoming more complex. The market is moving faster and the tools available to the manager are more sophisticated. This creates a need for "integrative models". These pull together diverse esults and relate the parts to the whole - in other words, translate everything into financial terms. The models discussed in this paper perform that role.

The simulation model has the potential to become the centerpiece of an overall portfolio management framework in the next decade. It will create a uniform environment for testing and refining policy initiatives - for example, marketing campaigns, collections actions, limit policies etc. The technology to deliver this complete solution is not yet mature. But there will be a significant competitive advantage to organizations which adopt this approach thoroughly and use it to develop a more coherent and rigorous approach to credit portfolio management.

The key concept was expressed by Einstein: "My goal is to make things as simple as possible - but no simpler". That is what bad debt modeling is all about.