MARKOV MODELS : AN INTRODUCTION
A New Approach to Bad Debt Modeling

Gerard Scallan

1. WHAT IS THE PROBLEM?

1.1 Every credit company needs to anticipate the payment behaviour of its clients. In particular, it needs to know the amount of write-off which will arise from the advances which it has already made. Bad Debt Forecasting is concerned to estimate:

a) The amount of currently outstanding debt which will not be repaid;
b) The timing of the write-offs;
c) If possible, the consequences for bad debt of decisions on the portfolio and trends in the credit market.

1.2 Bad debt forecasts usually take the form of provisions - percentages of the balances outstanding in various categories of debt are set aside to provide for future write-offs. In each accounting period, new provision is made based on the amounts entering each of these debt categories; amounts actually written off are deducted from the outstanding provision. The new provision is entered on the profit and loss statement. The net total provision is entered on the balance sheet.

1.3 Historically, provisioning levels have been established in a rough and ready manner. This is largely because:

a) the categories on which provision was made were subjectively defined;
b) the criteria for write-off were also subjective;
b) no computer records of accounts histories were available, so any estimation of percentages passing to write-off was based (at best) on small scale samples taken during audits.

1.4 This system of provisioning is still widely used, particularly on current account portfolios. It does not give rise to major problems where:

a) bad debt is insignificant in proportion to the overall portfolio;
b) the portfolio is stable, and hence bad debt levels are stable;
c) there are no marked trends in delinquency.

These conditions are increasingly rare; hence the historical method is increasingly insufficient.

1.5 When a portfolio grows rapidly, the historical method is likely to be very misleading. It normally makes provision only for late stages of delinquency. Therefore, no provision is made for new accounts going to write-off rapidly. Further, the categories are based on subjective judgment of accounts. The criteria are typically vague - for example "all reasonable measures for internal collection of the debt have been exhausted" - and describe the **actions** which have been taken on an account, rather than the **state** of the debt as such. When collection resources are under pressure, a backlog builds of accounts awaiting action. This means that the "action" criteria for passage to a different category cannot be fulfilled. But the lack of action is, in fact, aggravating the longer-term bad debt situation. Thus, the bad debt provision is of little value in accounting terms and is positively misleading as management information. (See the Monopolies and Mergers Commission Report on Credit Card Services, pp.186-191 for examples of fluctuating bad debt provisions).

1.6 The major alternative to the historical method is based on **roll-rates**. Accounts pass through several arrears stages - typically six to nine - before being written-off. A roll-rate is thought of as the probability of passing from one status to the next. The roll-rate from status k to status k+1 is defined as the amount in status k+1 in period n+1 divided by the amount in status k in period n. Table 1 gives an example.

Status	Balances - 31/05/90	Balances - 30/06/90	**Roll Rate**
Total Current	£25000k	£25200k	-----

Status	Balances - 31/05/90	Balances - 30/06/90	**Roll Rate**
Carried Fwd	????	£20000k	**2000/2500 = 80%**
New Spend	????	£5200k	-----
Arrears 1	£1500k	£1250k	**125/2500 = 5%**
Arrears 2	£400k	£300k	**30/150 = 20%**
Arrears 3	£240k	£240k	**24/40 = 60%**
Arrears 4	£200k	£180k	**18/24 = 75%**
Write-off	----	£160k	**16/20 = 80%**

Table 1 - Roll Rates Calculation

1.7 In order to calculate the future bad debt *arising from the amount currently outstanding*, we must take account of the payback rate. Therefore, a "current-current" roll rate is calculated. This is the ratio of debt outstanding from the previous period which is still outstanding - and in order -one period later. Thus, the write-off provision corresponds to what would arise if all spending were stopped on the portfolio and payments applied to existing debt.

1.8 The bad debt prediction is made by multiplying through the roll rates for a given initial distribution of debt. Table 2 applies the roll rates of Table 1 to the debt shown at end May. It is possible to calculate the write-off arising with an infinite horizon, by use of simp le formulae. A provision level corresponding to each arrears status can also be derived.

Status	Roll Rate	May	June	July	Aug.	Sep.	Oct.	Infinity
Current	80%	25000	20000	16000	12800	10240	8192	0
Arrears 1	5%	1500	1250	1000	800	640	512	0
Arrears 2	20%	400	300	250	200	160	128	0
Arrears 3	60%	240	240	180	150	120	96	0
Arrears 4	75%	200	180	180	135	112	90	0
Write-off	80%	0	160	154	154	108	90	0
Cumulative Write-off		**0**	**160**	**314**	**468**	**576**	**666**	**1006**
Outstanding		27340	21970	17610	14085	13360	9018	0

Table 2 - Application of Roll Rates

1.9 Implicitly, a roll rate model assumes that debt which does not "roll on" to the next state is repaid. This assumption is not correct. Usually, barely enough is repaid to restore the balance to order; the remaining balance is still at risk. This faulty assumption leads to systematic biases in the estimated provision levels and is especially prone to error in estimating the timing of write-off.

1.10 The roll rates model works tolerably for products where:

a) there is no spending except on accounts which are current;
b) delinquency is defined as time since last payment - i.e. in each period, all the accounts in a particular status either move on to the next state or return to current status;
c) all accounts in a given arrears status have about the same chance of rolling on the next status in the coming period.

1.11 Roll rate models assume that there is a strict succession of states, corresponding to months delinquent. The model has considerable difficulty where accounts can jump between arrears states, because of partial payments, where spending occurs on accounts in arrears or where accounts can stay in the same state for several successive periods. These can lead to roll-rates of more than 100% and cause severe anomalies in the results. All of these problems occur on current accounts. They lead to an erratic model, which normally underestimates the time before a write-off will show up but also underestimates the ultimate write-off arising from the outstanding balance.

1.12 Where a behavioural scoring system is in use, or there is some other statistical method to classify the risk on individual accounts, roll rates are unsatisfactory. They do not take this extra information into account.

1.13 Roll rate models do not allow for modeling of the effect of future spending patterns. Hence, they are of little help in the financial management of a portfolio.

1.14 The bad debt modeling problem is to find a methodology which:

a) allows delinquency states to be defined in more general terms;
b) which takes account of spending patterns on out of order accounts;
c) which is useful in making cash-flow projections;
d) which provides a natural way of exploring alternative scenarios of future payment patterns;
e) which can be adapted to deal with portfolios at various levels of complexity, depending on the purpose in hand.

2. MARKOV MODELS - THE IDEA

2.1 The solution to the problem lies in generalizing the roll rate approach. A powerful generalization can be obtained by using a statistical technique known as Markov Chains. Such models were first suggested for commercial lending portfolios in 1962 (Cyert et al. 1962). A more recent application is to be found in Corcoran, 1978. Further references are given in the bibliography. The ideas described here apply this technique to the consumer credit area and attempt to deal with the practical problems involved in modeling. This section describes the basics of how such a model works.

2.2 One of the main problems with roll rates is the assumption that accounts pass through a fixed sequence of states, neither stopping nor making partial recoveries, nor jumping states. Instead, we define a set of **"risk states"**. In general, an account can jump from any state to any other -or remain put - in successive periods. For instance, an account which was three months delinquent may make a partial payment and be classified as one month delinquent. On the other hand, if the customer's cheque bounces, the account will go to a high risk statestraight away.

2.3 Not all states need correspond to delinquency. For instance, the bad debt provision on an in-order account will vary considerably with its behavioural score. This can be accommodated by defining risk states which correspond to behavioural scorebands. In fact any group of accounts whose behaviour is of particular interest can be singled out as a risk state.

2.4 Table 3 gives an example of a set of states which could be used for current accounts - the type of account for which roll rates are least suited. We define two extra "fictitious" states - paid back and written off.

In-Order Balances	Mild out-of-order	Severe out-of-order	"Fictitious"
New a/c, low risk	<30 days, low risk	Arrangement, 3+ months	Paid back
New a/c, medium risk	<30 days, medium risk	Arrangement, 1-2 months	Written-off
New a/c, high risk	<30 days high risk, low balance	Promise to pay, no cheque book	
Primary a/c, never in debit	<30 days, high risk high balance	Abusive use w/ GCC	
Primary a/c, low risk	30-60 days, low balance, low risk	Default notice - no savings a/c	
Primary a/c, medium risk	30-60 days, low balance, high risk	Default notice - has savings a/c	
Primary a/c, high risk	30-60 days, high balance, low risk	Early Legal	
Secondary a/c, low risk	30-60 days, high balance, high risk	Late legal	
Secondary a/c, medium/high risk		Pending write-off	

Table 3 - Examples of Risk States, Current Accounts

2.5 On any given date, an account is classified in a single state. We can estimate the probability that an account moves from one state to another quite readily, using standard statistical techniques. However, bad debt forecasting is concerned with balances, not with numbers of accounts. Therefore, we need also to model the changes in balance over a period.

2.6 This is accomplished by treating the payment received as being separate from the state of the account. Suppose an account is in arrears with a debit balance of £1000 and makes a payment of £150 to clear the arrears. Of the £1000 in an arrears state at period 1, £150 moves to the Paid-back state, the remaining £850 moves to be a current balance in period 2. Therefore, the total balance is conserved. Thus, we use a form of "double entry book-keeping".

2.7 Balances can change for a number of reasons:

- payments and credits received reduce debit balances;
- interest increases balances by a fixed percentage;
- spending increases debit balances;
- fees are treated like spending.

In order to explain the overall changes in the portfolio, each of these changes needs to be accounted for.

2.8 Interest is charged on the balances in some states, but not in all. A spending rate can be calculated for each state, giving new spend as an absolute amount or as a percentage of balances in each state. Thus, we can reconcile the total balances outstanding in successive periods.

2.9 Table 4 illustrates how this analysis might work on the portfolio shown in Table 1. It assumes that accounts in areas 1 make a minimum payment of 10% to recover, those in arrears 2 make a 15% payment, in arrears 3 a 20% payment is needed and in arrears 4 a 25% payment is necessary. Spending is assumed to be proportional to balances for accounts which are in order or in Arrears 1. No spending is assumed in later stages of delinquency. Interest is charged up to arrears 4 at a rate of 2% per month.

June State	May State					Total Carried Forward	Interest	New Spend	June Balance
	In order	Arr. 1	Arr. 2	Arr. 3	Arr. 4				
Paid Back	3750	120	24	12	10	3916	---	---	---
In Order	20000	1080	136	48	30	21294	426	4911	26631
Arr. 1	1250	0	0	0	0	1250	25	288	1563
Arr. 2	0	300	0	0	0	300	6	0	306
Arr. 3	0	0	240	0	0	240	5	0	245
Arr. 4	0	0	0	180	0	180	4	0	184
Write Off	0	0	0	0	160	160			
Total	25000	1500	400	240	200	27340			
June Outstanding	21250	1380	376	228	30	23264	465	5200	28929

Table 4 - One-month Markov model

2.10 The key results are in the shaded areas. A total of £3916k in payments is received, but new spending amounts to £5200k. Thus there is a need to finance an extra £1284k - this is the net cash flow from the portfolio. £160k is written off during the course of the month. The last column shows the resulting portfolio structure in June. The total amount outstanding is £28929k, of which £26631k is current. This takes account of interest payments, of new spending and of the progression of accounts through various areas states. Thus it is possible to get a complete picture of the state of the portfolio, month by month - cash flow, outstanding balances and risk structure.

2.11 In practice, fewer of the cells are zero; some accounts will make partial payments and move from Arrears 4 to Arrears 2; others may declare bankruptcy and move from current to write-off in a single period. "Generalised roll rates" are calculated, which give the proportion of the balance outstanding in one state which moves to any other in the next period. This overcomes the restriction on the nature of the states.

2.12 These generalised roll rates are combined in a set of mathematical equations known as a Markov model. This allows the state of the portfolio to be projected forward one period at a time. By compounding these one period projections (as with roll rates), projections can be made over longer periods. If quarterly periods are used, projections over a two to three year horizon are likely to be quite reliable. As with any statistical model, longer horizons must be treated prudently. The model is useful in longer term planning. However, a variety of scenarios for trends in the generalised roll rates should be used. This tests the realism of the assumptions made.

2.13 The model suggested here is very general. It can be used to project bad debt and to assign provisioning levels. However, it can also generate funding requirements and be used to project portfolio growth. It also allows for "what if" simulations of the effect of changing various policies - not just in relation to bad debt, but also in relation to recruitment, to interest rates, to limit policies and to collections effort.

3. DEVELOPING THE MODEL

3.1 The development of a bad debt model involves two phases - an initial set-up followed by a continued checking, maintenance and updating. It is extremely important to pick up trends early, yet not to be misled by "statistical flukes".

3.2 The first step in developing a Markov Bad Debt Model is to assemble a database. For each account to be used, summary information is required to cover at least twelve months (providing the account has been open that long). By preference, two years information should be used. This will allow investigation of seasonality not possible in a single year.

3.3 The information required for each account in each period is:

- opening balance;
- total credits received (broken down between payments, credit interest and other credits if possible);
- total expenditure;
- debit interest;
- closing balance;
- information sufficient to define the likely states of interest.

3.4 The information required to define states of interest will vary considerably from one portfolio to another. In large measure, the definitions depend on the available information. For instance, if there is no behavioural score, it is pointless to have states depending on scoreband. The following list shows fields likely to be of use:

- current delinquency (or days dormant);
- days in debit during period (current accounts only);
- time on books;
- type of account (e.g. budget/option);
- time since last delinquent;
- behavioural score;
- account blocks;
- turnover of account.

3.5 The information is required for a large sample. If possible, the whole portfolio should be used. The large sample is needed because small proportions of accounts go to write-off in any period. Hence, to reliably estimate the proportions. In the set-up phase, it is possible to use a more modest sample (with richer information) to identify the states to be used. Then a larger sample will be taken, stratified by state, to estimate the generalised roll rates. In the maintenance phase, it is not possible to stratify based on outcome state. In this situation, a sample of about 1000 accounts in a particular state is needed to estimate the generalised roll rates sufficiently accurately. As with any statistical system, a validation sample should be constructed.

3.6 Once the sample has been assembled, an initial (small) set of states is defined. Then a cluster analysis is performed to identify families of accounts which are relatively homogeneous in terms of their generalised roll rates. These families then become the states in a more elaborate system.

3.7 Time series methods are applied to estimate the generalised roll rates (See section 5). Then, statistical tests check that the states really do capture all the essential information about the roll rates - that is, all accounts in a given state have roughly the same probability of jumping to each other state next period. If not, the states are revised and the process starts over.

3.8 Once a satisfactory set of states and of generalised roll rates has been identified, the system is validated on a hold-out sample.

3.9 Separate systems are possible for different subpopulations. However, within an overall system, certain subpopulations (such as those based on time on books) will tend to move within separate groups of states. Thus, separate systems are not usually needed.

3.10 Once the system has been set-up, it should be tested each period. This test consists of taking the predictions based on the previous periods state distribution and verifying the accuracy (or otherwise) of its predictions. If major discrepancies occur, then one of two things is happening:

- roll rates are changing due to general trends (e.g. change in economic conditions); or

- the state structure has become unstable (due to underlying shifts in the behavioural patterns on the portfolio).

In the first case, experimentation will help determine the likely outcome should the trend continue. In the second, the least stable states should be examined with a view to restructuring - merging or splitting states. This is of lesser concern from a bad debt viewpoint.

3.11 Approximately once a year, all the states should be examined to see if they are still reasonably homogeneous. This is likely to lead to the redefinition of some states.

4. SYSTEM USE

4.1 Once a period, an update of the Model Database is made. It adds the details of each accounts behaviour for the period just ended. Using the previously existing generalised roll rates, new estimates are prepared. These start with the existing distribution of outstanding balances over states, and project how this will move between states in the following periods.

4.2 Apart from the base projection, alternatives will show the effect on portfolio structure of deterioration or improvement in the generalised roll rates. These estimates will serve to give management an indication of the potential effect of changes in underlying trends, due to changes in the economy or other factors external to the model.

4.3 The model can be projected forward as far as is wished. However, the accuracy of the results is obviously less over longer time periods than over shorter ones. The variations on the basic set of generalised roll rates will serve to give management an idea of the accuracy of the longer term estimates. An "ultimate" provision estimate can also be derived - of the amount of debt in each state, what proportion will ultimately be written off. Discount factors can be applied to calculate the net present value of the debt in various states.

4.4 The underlying model will run on a mainframe computer. However, summaries, including the amount of debt in each state, and basic variants on the roll rate matrix can be down loaded to a Personal Computer. Here, their structure is ideally suited to manipulation within a spreadsheet. Thus, managers can formulate their own hypotheses about future trends. This kind of model is am important component of Executive Information Systems of the future. Graphic presentations will allow managers to grasp more readily the significance of the numbers generated.

5. THE MATHEMATICS

5.1 Let $\mathbf{x}(t)$ be an N+1 dimensional vector, representing the state of the portfolio at any particular period, t. $\mathbf{x}_0(t)$ is the amount of debt paid back up to period t, set to zero at the outset. States 1 to N-1 are the regular states of the portfolio. State N corresponds to written off debt.

5.2 Let $\mathbf{A}$ be an (N+1)x(N+1) dimensional matrix, which gives the transition probabilities from one state to another. That is $\mathbf{A}_{ij}$ is the proportion of the debt in state i at period t which is in state j at period t+1. It is assumed that $\mathbf{A}$ is stationary - i.e. that the transitions do not depend on t. (This assumption is tested later on). $\mathbf{A}$ is called the roll-rate matrix.

5.3 Let $\mathbf{S}$ be a diagonal matrix of order N+1. The diagonal elements of $\mathbf{S}$ give the propensity to spend of accounts in each state; that is $\mathbf{S}_{ii}$ is the ratio of amount spent to balance in period t by accounts in state i *during period t+1*. Obviously, no spend is allowed in states 0 or N. $\mathbf{S}$ is called the spend matrix.

5.4 Let $\mathbf{R}$ be a diagonal matrix of order N+1. The diagonal elements of $\mathbf{R}$ give the period interest rate to be applied to accounts in state i. $\mathbf{R}$ is called the Interest matrix.

5.5 The one-period model is:

$$\mathbf{x}(t+1) := (\mathbf{I} + \mathbf{R})(\mathbf{I} + \mathbf{S})\mathbf{A}'\mathbf{x}(t) \qquad t = 0,1,2,....$$

5.6 The appropriate level of provision for each state is given by:

$$\mathbf{r} := (\mathbf{I} - \mathbf{A}_{\{1..N-1\}})^{-1}\,\mathbf{a}_{\{1..N-1\},N}$$

where $\mathbf{r}$ is the provision level for each state (a vector of order N-1), $\mathbf{A}_{\{1..N-1\}}$ is the vector of transition probabilities restricted to states 1..N-1 and $\mathbf{a}_{\{1..N-1\},N}$ is the vector of direct transition probabilities for states 1..N-1 to state N (write-off).

5.7 The expected cash flow under any scenario is given by

$$\lim_{t\to\infty} \mathbf{x}_0(t)$$

5.8 The expected write-off is given by

$$\lim_{t\to\infty} \mathbf{x}_N(t)$$

5.9 The model can be discounted by an appropriate factor by including a discount matrix $\mathbf{D}$ in equation 5.5:

$$\mathbf{x}(t+1) := (\mathbf{I} - \mathbf{D})(\mathbf{I} + \mathbf{R})(\mathbf{I} + \mathbf{S})\mathbf{A}'\mathbf{x}(t) \qquad t = 0,1,2,....$$

$\mathbf{D}$ is a diagonal matrix with $\mathbf{D}_{00} := 1$, $\mathbf{D}_{ii} := 1\text{-}d$ for i = 1,2,..N-1, $\mathbf{D}_{NN} := 1$, where d is the period discount factor to be applied. Note that only the "live" states (1..N-1) are discounted. States 0 and N are accumulations of values that enter at different periods. Hence, the values must be discounted before they enter, not after.

5.10 The statistical problems are:

a) estimating $\mathbf{A}$ and $\mathbf{S}$;
b) determining a sufficient state space, so that the process determined by $\mathbf{A}$ is in fact Markovian.

5.11 An appropriate way to address the problem is to treat $\{\mathbf{x}(t)\}$ as a multi-dimensional auto-regressive process. Certain coefficients are constrained to be zero. The coefficients can be estimated using the Box-Jenkins approach. We test the hypothesis that the process is Markovian by the testing the order of the autoregressive function. If it is one, then we conclude that the model is Markovian. An alternative best fit is constructed using Ake-Ike's Information Criterion.

5.12 Alternatively, the matrix $\mathbf{A}$ can be estimated by using a jack-knife approach. The variance of estimate can then be found using a bootstrap method. This has the advantage that the constraints on the coefficients (zeroes and stochastic condition) will be met automatically. The Markovian condition is then checked by testing for equality between $\mathbf{A}^2$ and an $\mathbf{A}$-type matrix constructed over two periods.

5.13 The principal theoretical difficulties with this approach are:

a) finding the appropriate transformation to turn $\{\mathbf{x}(t)\}$ into a stationary time series;
b) dealing with seasonality.

5.14 Tests for trends in generalised roll rates are made by testing the sample value of the autoregressive function against the historically derived estimate.

6. BIBLIOGRAPHY

Bierman,H., Hausman,W.H. (1970): "The Credit Granting Decision", Management Science, 16/8, pp. B519-532, April 1970.

Bunn,D.W. (1981): "An Empirical Bayes Procedure for the Credit Granting Decision", Ops.Res. Letters **1**, pp. 10-12, October 1981.

Corcoran, A.W.(1978): "The Use of Exponentially Smoothed Transaction Matrices to Improve Forecasting of Cash Flows from Accounts Receivable", Management Science, **24**, 7 (March 1978).

Cox,J., Stanford,R.E., Dean,E. (1978): "An Investigation of the Use of Markov Analysis in Projecting Short-term Cash Needs and Bad Debts in the Banking Industry", Auburn University School of Business Working Paper, 1978.

Cyert,R.M., Davidson,H.J., Thompson,G.L. (1962): "Estimation of the Allowance for Doubtful Accounts by Markov Chains", Management Sci., **8**, March 1962.

Cyert,R.M., Thompson,G.L. (1968): "Selecting a Portfolio of Credit Risks by Markov Chain", Journal of Business, pp. 39-46, January 1968.

Dirickx,Y.M.I., Wakeman,L. (1976): "An Extension of the Bierman-Hausman Model for Credit Granting", Management Sci., **22**, pp. 1229-1237, July 1976.

Frydman,H., Kallberg,J.G., Kao,D-L. (1985): "Testing the Adequacy of Markov Chain and Mover-Stayer Models as Representations of Credit Behavior", Ops.Res.,**33/6**, pp. 1203-1214, November 1985.

Liebman, L.H. (1972): "A Markov Decision Model for Selecting Optimal Credit Control Policies", Management Sci.,**18**, pp. B519-525, June 1972.

Stanford,R.E. (1989): "Optimizing Profits from a System of Accounts Receivable", Management Sci.,**35/10**, pp.1227-1235, October 1989.

Srinivisan,V., Yong,K.H. (1987): "The Bierman-Hausman Credit Granting Model: A Note", Management Science, **33/10**, pp. 1361-1362, October 1987.

Thomas,L.C. (1988):"Behavioural Scoring", Working Paper, University of Edinburgh.