Time will tell: Behavioural scoring and the dynamics of consumer credit assessment

L.C.Thomas*, J.Ho** and W.T.Scherer***,

* School of Management
University of Southampton
Southampton
SO17 1BJ
U.K.

** Department of Business Studies,
University of Edinburgh,
Edinburgh,
U.K.

*** Department of Systems Engineering
University of Virginia
Charlottesville, VA 22903
U.S.A.

Abstract

This paper discusses the use of dynamic modelling in consumer credit risk assessment. It surveys the approaches and objectives of behavioural scoring, customer scoring and profit scoring. It then investigates how Markov chain stochastic processes can be used to model the dynamics of the delinquency status and behavioural scores of consumers. It discusses the use of segmentation, mover-stayer models and the use of second and third order models to improve the fit of such models. The alternative survival analysis proportional hazards approach to estimating when default occurs is considered. Comparisons are made between the ways credit risk is modelled in consumer lending and corporate lending.

Keywords: behavioural scoring, Markov chains, survival analysis, credit risk modelling.

1. Introduction

Application scoring (see Hand, 2001, in this issue for a review) in consumer credit risk assessment consists of connecting two snapshots of the state of the consumer – the first of their characteristics on application and the second of their creditworthiness at some later date. Thus it is a static phenomenon. Behavioural scoring on the other hand is a way of updating the assessment of consumer credit risk in the light of the current and most recent performance of the consumer. Thus it replaces the first snapshot by a description of the dynamics of the consumer's recent performance, but the second snapshot still remains.

When one considers the profitability either of a customer on a specific product or of the total profitability of a customer to a lender, one needs to use the recent consumer behaviour to estimate subsequent performance over a future time interval, not just at some specific future time. Thus, to develop customer profit scores one needs to estimate the future dynamical behaviour of the consumer. One needs a forecast of the dynamic behaviour of the behaviour score itself or the delinquency status of the consumer. The latter would be a way of estimating how much default there will be in each subsequent period for a given portfolio of consumer loans. Such calculations are needed to forecast how much the lender needs to put aside to cover these expected losses – the debt-provisioning problem. We investigate how Markov type probability models could be used to obtain this estimate.

One can use models based on survival analysis ideas to estimate when customers will default. Such models allow one to estimate the profitability of customers on a product, since they can deal not just with default risk but also other profit lessening events, like early repayment of a loan. These approaches connect the recent dynamical behaviour of a consumer with the dynamical behaviour of the probability of default and other measures of consumer behaviour over the whole future. Thus one has transformed both snapshots of application scoring into "movie clips" of the consumer's behaviour.

In section two we describe the difference between behavioural scoring and application scoring and review the types of decisions that the former is used to make. Behavioural scoring has been in operation since the late 1960s, when Fair Isaac Inc. introduced such a system for Montgomery Ward (Lewis, 1992). A detailed account of how such systems are used in practice is given in Hopper and Lewis (1992) and in Chapter 7 of McNab and Wynn (2000). Most such behavioural scoring systems are statistically based, but there are a number of probability based behavioural scoring models that have been suggested, based on the original ideas of describing the consumer's behaviour by a Markov chain (Cyert, Davidson and Thompson, 1962). These have been reviewed in Thomas (2000).

Section three looks at how one can change the objective in behavioural scoring from estimating default risk to estimating either the profit on the product or the total customer profit. This idea of combining risk and return was suggested by Hoadley and Oliver (1998) and the problems in scoring the whole customer are alluded to in the reviews by McNab and Wynn (2000) and Thomas (2000).

The Markov chain approaches to modelling the dynamics of consumer behaviour may not have become the industry norm for behavioural scoring, but they have found favour in estimating the probability to default (PTD) needed for debt provisioning. Markov chain models have been used in a number of different contexts in the last two decades – including road maintenance (Golabi, Kalkarni and Way, 1982), bridge repair (Scherer and Glagola,

1994) and health care (Fuller and Scherer, 1999). As the work by Weiss, Cohen and Hershey (1982) on hospital patient flow suggests, the difficulty in using such models is segmenting the population and then choosing appropriate state classifications for the different segments, so that the resulting flow is Markov or almost Markov. Section four looks at the Markov chain approach to behavioural scoring, while section five outlines how such models can describe the dynamics of consumer repayment behaviour. Accurate models require great care in segmenting the population into subpopulations and defining the states for each segment so as to ensure Markovity.

Section six provides a brief outline of the survival analysis approach to estimating not if but when consumers will default. This approach was first suggested by Narain (1992) and has been developed recently by Stepanova and Thomas (1999, 2001) and Hand and Kelley (2000).

In the conclusion, common features of and differences between the models used in estimating the dynamics of credit risk in consumer lending and those used in estimating the dynamics of credit risk in corporate lending are identified.

2. Behavioural Scoring

Behavioural scoring uses characteristics of customers' recent behaviour to predict whether or not they are likely to default. The methodology is very similar to that of application scoring. A sample of customers is chosen so that the data on their transaction performance either side of an arbitrarily chosen observation point are available. The period before the observation time is called the performance or observation period and is usually 6 to 12 months. The characteristics that will be used in the behavioural scorecard describe the customers' performance during this time. Typical variables would be average, maximum and minimum levels of balance, credit turnover, and debit turnover. Other characteristics estimate the trend in payments or balances during the period either by taking weighted averages or taking ratios of performance in the latter part of the period compared with that in the earlier part. Some of the characteristics are indicators of delinquent behaviour – number of missed payments or times over overdraft or credit limit - while others reflect difficulty in money management such as the number of cash advances using a credit card. A pure behavioural scoring system will only include variables dealing with the customers' performance and the current values of variables from monthly credit bureau reports. Other behavioural systems include personal characteristics such as age, time with bank or residential status as well the pure behavioural characteristics.

The period after the observation point is the outcome period, which is often taken as 12 months, and the customer is classified as a good or a bad depending on their status at the end of this outcome period. A common definition is to classify a bad to be someone who is 90 days overdue at this point. It is not the case that all other customers are classified as good. In order to separate the goods and the bads as much as possible, those with behaviour that is not yet bad but is tending that way are classified as indeterminate and left out of the sample. Thus those between 30 and 90 days overdue may be put in this category and the goods are then those who repayments are up to date or, at most, less than 30 days delinquent.

The methodologies described in Hand (2001) are then used to build a scorecard that best classifies the goods and the bads. One important consideration is whether to segment the population and build different scorecards on each segment. There are three reasons for segmenting scorecards – strategic, operational and variable interactions. Some banks may decide to target certain groups of customers, depending for example on their age or their residential status. They prefer to have a separate scorecard for these groups because they may wish to treat them differently in the future, by taking a greater risk exposure with them and so having a lower cut-off score. New customers with little credit history must have a separate scorecard because the characteristics in the standard scorecard do not make sense operationally for them. Similarly, customers who have no borrowing facility cannot become delinquent if they subsequently borrow and so may need a separate scorecard that does not involve delinquency characteristics. Finally there may be strong interactions between important variables. If the interaction is only between one pair of variables it may be sufficient to include the combined variable in the scorecard. If, however, one characteristic interacts strongly with a number of others then it may be sensible to segment the population according to attributes under this characteristic.

One of the disadvantages of behavioural scoring is that one typically needs two years history to build a scorecard and thus the population one then applies it to may be quite different from that it was built on. One way used to cut this down (as well as taking performance periods of only six months) is to take a shorter observation period - say six months - and classify customers as bad if they exhibit characteristics at the end of this period that suggest they may subsequently go bad. These characteristics can be obtained by building a separate scorecard on a different sample to find which characteristics are indicative that the customer will go bad within a further six months.

This lag, between the period of time when the transaction information that was used to build the scorecard was collected and the period of time when the scorecard is used, means that both the population characteristics and the economic environment may have changed. The latter problem is heightened because behavioural scorecards tend to have no external economic characteristics in them. The unwritten assumption is that the relationship between the performance characteristics and the subsequent delinquency status of a customer will be the same now as it was two to three year ago when the information on which the scorecard was built was collected. This is assumed to be the case no matter what economic changes have occurred in that period.

Hopper and Lewis (1992) and McNab and Wynn (2000) both give accounts of how behavioural scoring systems can be used in practice. As well as setting credit limits, authorizing accounts to go into excess and pre-authorization of direct mailing offers, behavioural scoring can be an input into deciding how to deal with those in arrears. They advocate experimentation using a champion/challenger approach. In this, one splits the customers randomly and applies different collection policies to each to find out which works best on which cohort of customers, grouped according to behavioural scores and other characteristics. One uses the existing the policy (champion) for the majority of the customers and tries the new policy (the challenger) on a much smaller subset until it is clear which is more successful.

3. Profit and customer scoring

Behavioural scorecards have typically been applied to the customers for one loan product using their behaviour on that product. This is an example of product default scoring. More recently it has been realised that customer performance on one product may give good indications of their likelihood to default on other products. In particular, if a bank has a customer's main current account or cheque account, it is a very good indicator of the general economic health of the customer. Changes in behaviour in that account may well presage delinquency in loan accounts. Thus scorecards have been developed using characteristics on all the customer's products with the lender to try and estimate the chance of defaulting on all or some of the loans. This is referred to as customer scoring or, more properly, customer default scoring (see McNab and Wynn, 2000) and the methodology is that of standard behavioural scoring.

The competition in the lending market has made lenders think about the profitability of a loan as well as its default risk. Ideally a bank would like to score the profitability of giving that customer that particular credit line – a product profit score. Even more useful would be to develop a scorecard that assesses the profitability of the customer to the lender over all products – a customer profit score. Some progress has been made in this direction, but as Thomas (2000) points out, a real profit scoring system would need to develop new approaches to modelling consumers' performance. This is because to measure the profitability of a customer one needs to record their behaviour over a suitable time interval – not just record their status at one time point - which is the nub of default scoring. Thus one needs to model the dynamics of the customer's behaviour. Two such models – Markov chains and survival analysis - are outlined in this paper.

The only approaches to profit scoring that have been implemented commercially to date are to band customers according to a risk measure and a return measure and apply different policies to each joint band. For example, some lenders set overdraft limits by constructing a matrix of bands of behavioural scores (risk) and of average balance or some more sophisticated measure of return, as in Table 1. Judgement is used to set the overdraft limits for each cell of the table.

Thus despite the sophisticated modelling of the default risk, there is no real modelling of total profit nor of the way the decisions made affect the profit.

Table 1: Overdraft limit as a matrix of risk and return

Overdraft Limit	Balance< £500	£500 < Balance < £2500	Balance > £2500
Beh. Score > 500	£10,000	£12,500	£15,000
300 < Beh. Score < 500	£2,000	£4,000	£10,000
Beh. Score < 300	No overdraft	£500	£1,000

4. Markov chain based models

Markov chain based models of consumer behaviour provide an alternative approach to behavioural scoring and have obvious extensions to profit scoring. These models were first suggested by Cyert, Davidson and Thompson (1962) and variants of the basic model were suggested by Bierman and Hausman (1970), Corcoran (1978) and van Kuelen, Spronk and Corcoran (1981). However, there have been very few commercial systems based on these ideas; yet by extending the ideas from Markov chain models to Markov decision process models (Thomas, 1994) one can build profit-scoring systems that give model-based decisions on overdraft limits rather than the subjective ones described above.

An example of such a model is as follows. The state, u, of a customer's account is given by a triple u = (b,n,i) where b is the balance outstanding on the account, n is the number of periods since last payment, and i describes any other relevant information. The decision to make is "what is the credit limit, L, in each of these states?". To do this one needs to estimate $p^L(u;u')$, the probability that the account goes from u to u' under credit limit L. One also needs to calculate $r^L(u)$, the profit to the lender if the customer is in state u and credit limit L is applied.

$p^L(u;u')$ is obtained by estimating:

$t^L(u;a)$, the probability an account in state u with credit limit L repays a next period;

$q^L(u;o)$, the probability an account in state u with credit limit L orders o next period; and

$w^L(u;i')$, the probability an account in state u with credit limit L changes its information to i'.

These probabilities can be obtained empirically by estimating the transitions in the histories of a sample of customers. One can then define the transition probabilities from one state to another by summing terms corresponding to probabilities of outcomes where there is both payment and purchase, payment only, purchase only neither purchase nor payment and which give rise to the required changes in b,n and i. The probabilities in these four types of outcome are defined as follows,

$$p^L(b,n,i;\ b+o-a,0,i') = t^L(u;a)\, q^L(u;o)\, w^L(u;i'), \quad \text{provided } b+o-a \le L, \text{ and } a>0; \qquad (1)$$

$$p^L(b,n,i;\ b-a,0,i') = t^L(u;a)\, w^L(u;i')\, \big(q^L(u;0) + \textstyle\sum_{o>L-b+a} q^L(u;o)\big), \quad \text{where } a>0; \qquad (2)$$

$$p^L(b,n,i;\ b+o,n+1,i') = t^L(u;0)\, q^L(u;o)\, w^L(u;i'), \quad \text{provided } b+o \le L; \qquad (3)$$

$$p^L(b,n,i;\ b,n+1,i') = t^L(u;0)\, w^L(u;i')\, \big(q^L(u;0) + \textstyle\sum_{o>L-b+a} q^L(u;o)\big). \qquad (4)$$

These expressions depend on independence between the purchase, repayment and status processes – an assumption implicit in the way the probabilities t, q and w were estimated.

Note that in (2) and (4) there may be no new order either because none was made or because the new order would take the customer over their credit limit.

If one assumes that a fraction, f, of the purchase price is profit (including any profit from interest on the repayments), and that the lender writes off the bad debt after N periods of non-payment, the profit in any one period is then

$$r^{L}(b,n,i) = f\,(\sum o\, q^{L}(u,o) - b\, t^{L}(u,0)\, \delta\, (n-(N-1)) \tag{5}$$

The second term reflects the assumption that a customer is considered to be in default when no payment is made for N periods. One can then apply the standard dynamic programming approach and show that $V_n(u)$, the expected profit over n periods given account in state u, satisfies the optimality equation

$$V_n(u) = \max_L \left\{ r^L(u) + \sum_{u'} p^L(u,u') V_{n-1}(u') \right\} \tag{6}$$

Solving this would give the credit limit that maximises the profit over n periods.

This uses an orthodox statistical approach in that the parameters of the transition matrix are estimated from past data on other customers. Bierman and Hausman (1970) suggested that these parameters could be estimated in a Bayesian way, with the belief about the parameters of each customer being updated in the light of their own payment performance.

5. Modelling the dynamics of behavioural scoring and delinquency

The Markov chain model of consumer behaviour depends on two crucial assumptions. First, that the state space of the model does describe all the different situations that the consumer can be in, and second, that the dynamics of their subsequent behaviour does follow Markov behaviour. It is this latter assumption, that there is a simple stochastic model of the dynamics, which allows one to calculate the expected future profitability of each customer. One could hope that the same type of probabilistic modelling of the dynamics would work on other aspects of consumer behaviour, including both their delinquency status and their behavioural score.

Although Markov chain models are not widely used to build behavioural or profit scoring systems, they are used widely to describe the dynamics of the delinquency status of a population. This can be used to estimate the expected loss due to default in the portfolio in future time periods and hence is an aid to debt provisioning. Alternatively, the estimates of the numbers of delinquents and defaulters in different time periods can be used to plan the resources needed in the collections and recovery departments.

The models in use at present are fairly straightforward. The states are the different delinquency states – say 0,1,2,3,4+ months past due. The transition probabilities or the roll rates are obtained from past data. Take a sample of customers and assume their dynamical performance is stationary. Let n(i) be the total number of months customers are in state i (i=0,1,2,3,4) and let n(i,j) be the number of times that customers move from state i to state j. Bartlett (1951) and Hoel (1954) have shown that the maximum likelihood estimate of the transition probability p(i,j) is n(i,j)/n(i). Thus in table 2, if the upper number gives the number of such transitions in the sample, the lower number gives the maximum likelihood estimate of the transition probabilities.

Next Current	0	1	2	3	4
0	19700 0.985	300 0.015			
1	100 0.25	160 0.4	140 0.35		
2	7 0.047	8 0.053	45 0.3	90 0.6	
3	5 0.05	1 0.01	4 0.04	15 0.15	75 0.75

Table 2: calculation of transition probabilities

This approach allows the data to define the transition matrices, but it may be sensible to put some restrictions on this. Thus certain transitions may be deemed impossible. This would introduce structural zeros into the matrix and has the advantage of limiting the number of parameters that have to be estimated. In Table 2, one might say that the transitions $0 \rightarrow 2$, $0 \rightarrow 3$, $0 \rightarrow 4$, $1 \rightarrow 3$, $1 \rightarrow 4$, and $2 \rightarrow 4$ are not possible and that one may assume $3 \rightarrow 1$ is so unlikely as to be ignored.

Having calculated the transition probability matrix P and given $\pi(0)$ the current distribution of the population between the states, then the expected distribution in m periods time will be $\pi(m)= \pi(0)P^m$. One has to modify this calculation to allow for attrition - customers who finish their association with the lender - and for new customers arriving. Thus one has to be careful to make sure whether one is calculating the delinquency status of the cohort who were customers as of time 0 or the delinquency status of the current population. The latter is constructed by adding together cohorts each consisting of customers who joined the lender in the same time period.

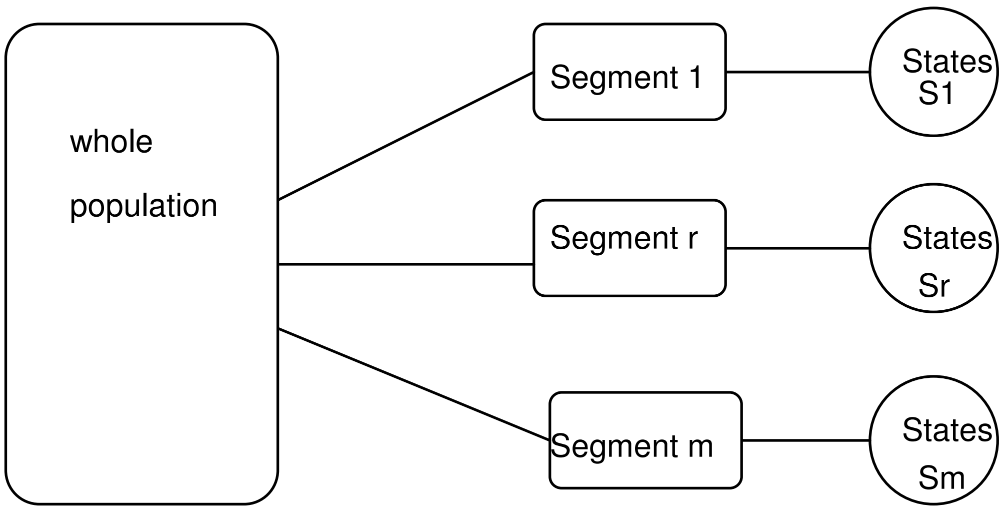

Figure 1: Segmentation of population

One needs to be confident that the dynamics of the model reflects the reality of the dynamics of the population. It is almost never the case that all customers will follow the same stationary Markov process. So the problem is to define a set of subpopulations $r \in R$ and sets of states, S^r , for each such subpopulation, r, so that the process for each subpopulation is Markov, see Figure 1. In the delinquency models, the initial choice of states will involve conditions on the numbers of days past due together with conditions on the amount of the excess, to avoid insignificant debts being considered. In a behavioural scoring model, the states will be bands of the behavioural score.

As in behavioural scoring one cannot easily separate the segmentation process from the choice of states in each segment, (though here one is segmenting to improve the dynamics of the model rather than its classification accuracy). Since one is seeking processes that are as nearly Markov as possible, one of the most useful tools is the χ^2 tests for Markovity, first suggested by Anderson and Goodman (1957). The idea is to compare the frequency with which the sequence of state transitions $a \rightarrow j \rightarrow k$ occurred compared with $b \rightarrow j \rightarrow k$ for all k. If the process were truly Markov then these distributions should be the same for all choices of a and b.

Segmentation into subpopulations is done for three reasons. One may use intuition and segment by the mix of financial products being held by the consumer. If the lender holds the consumer's main current account there is much more information available to model the consumer's situation than if that account is not available. Mortgage accounts perform differently from personal loans. A second type of segmentation is by the age of the account. Consumers who have an established history with a lender are generally more stable than those who have only recently opened borrowing facilities, simply because the more volatile of their vintage have defaulted or moved to other lenders. The third reason to segment is because of the behaviour of the account itself. One wants segments each of which is homogeneous in terms of its behaviour. One split that appears to do this quite well is the mover-stayer model. The idea of mover-stayer appeared first in labour mobility studies and subsequently was used in consumer purchasing behaviour. Frydman, Kallberg and Kao (1985) were the first to suggest its use in the consumer credit context and developed estimators for the parameters required (Frydman, 1984). Related estimates were developed by Weiss, Cohen and Hershey (1982). In the context of consumer credit, stayers are those who pay off their debt fully each month and so always remain in the highest good state. Movers are customers whose payment history is more varied, including partial and missed payments. Some detailed analysis of these concepts in the case of a large bank's customer base (Ho, 2001) suggested the split between the two groups is about 50:50.

Even with segmentation it is likely that models built on the initial choice of states are far from Markov. The χ^2 values in the Anderson-Goodman Markovity test will be way above the range for accepting the null hypothesis. In such cases, it is necessary to see whether more complex state definitions will preserve Markovity. In particular if one defines a second order Markov chain, so that the "state" at any time is the current basic state and the basic state at the previous period. This increases the number of states considerably but many of the transitions are now not possible. However, it is surprising how often this second order state system is almost Markov. This is what Golabi, Kalkarnia and Kao (1985) found in their road maintenance models (though looking at the model after ten years of operation, Wang, Zaniewski and Way (1994) believed a first order chain would be sufficient). Fuller and Scherer (1999) found a second order chain modelled the situation well in their work on healthcare expenses. If even this is not satisfactory, it may be necessary for some segments of the population to go to a third order Markov chain, where the "state" is the current and the previous two basic states the customer has been in. This is very likely to satisfy the Markov requirement, but the matrix itself is extremely sparse. If there were N original basic states then only $1/N^2$ of the transition matrix entries will be non-zero. However for some very volatile segments it has been necessary to model at this level of complexity.

Even when Markovity has been achieved by segmentation and careful state definition, the resultant processes may well be non-stationary, as the transition probabilities are likely to depend on

- the age of the accounts, s,

- the time period, t and
- external economic effects, like the interest base rate, e.

So one tries to estimate transition probabilities $p^r_{jk}(s,t,e)$, which is the probability of a customer in subpopulation r moving from state j to state k, in period t, when their account is aged s, and the current base rate is i. One model that has been implemented (by Ho, 2001) was to segment by age of account and for each segment define the transition probabilities $p_{jk}(t,e)$ for $0 \le t \le T$ and $0 \le e \le E$ by

$$p_{jk}(t,e) = p^0_{jk} + a_{jk}t + b_{jk}e \qquad (7)$$

$$\text{with } \sum_k p^0_{jk} = 1;\ \sum_k a_{jk} = 0; \sum_k b_{jk} = 0; p^0_{jk} \ge 0 ;\ p^0_{jk} + a_{jk}T + b_{jk}E \ge 0$$

This gave a good fit with reality and the signs of the a's and b's made sense in terms of the factors affecting delinquency.

6. Survival analysis approach to profit scoring

The Markov chain models describe the dynamics of a consumer's movement through a number of delinquency states or scoring bands. If one is only interested in when they reach the default state and not their intermediate behaviour then one can use survival analysis approaches to estimate when this will occur. So instead of just asking which consumers will default, as in behavioural scoring, one asks when will they default

Using survival analysis to answer the "when" question has several advantages, namely:

i. it deals easily with censored data, where customers cease to be borrowers (either by paying back the loan, death, changing lender) before they default;

ii. it avoids the instability caused by having to choose a fixed period to measure satisfactory performance, which is inherent in behavioural and application scoring;

iii. estimating when default occurs is a major step towards calculating the profitability of an applicant; and

iv. it makes it easier to incorporate estimates of changes in the economic climate into the 'scoring' system.

Narain (1992) was one of the first to suggest that survival analysis could be used in credit scoring. Banasik, Crook and Thomas (1999) compared the survival analysis approach with logistic regression based scorecards and showed how competing risks can be used in the credit scoring context. Stepanova and Thomas (1999,2001) and Hand and Kelley (2000) developed the ideas further and introduced tools for building survival analysis scorecards, as well as introducing survival analysis ideas into behavioural scoring.

If T is the time until a loan defaults then there are three standard ways of describing the randomness of T in survival analysis:

survival function $S(t) = \text{Prob}\{T \geq t\}$ where $F(t) = 1-S(t)$ is the distribution function;

density function $f(t)$ where $\text{Prob}\{t \leq T \leq t+\delta t\} = f(t)\delta t$; and

hazard function $h(t) = f(t)/S(t)$ so $h(t)\delta t = \text{Prob}\{t \leq T \leq t+\delta t | T \geq t\}$.

In the survival analysis approach, we want models that allow the application and behavioural characteristics to affect the probability of when a customer defaults. Two models connect the explanatory variables to failure times in survival analysis – proportional hazard models and accelerated life models. If $\mathbf{x} = (x_1, \ldots x_p)$ are the explanatory characteristics, then an accelerated life model assumes

$$S(t) = S_0(e^{\mathbf{w.x}} t) \text{ or } h(t) = e^{\mathbf{w.x}} h_0(e^{\mathbf{w.x}} t) \tag{8}$$

where h_0 and S_0 are baseline functions so the **x** can speed up or slow down the 'ageing' of the account. The proportional hazard models assume

$$h(t) = e^{\mathbf{w.x}} h_0(t) \tag{9}$$

so the characteristics **x** have a multiplier effect on the baseline hazard. One can use a parametric approach to both the proportional hazards and accelerated life models by assuming $h_0(.)$ belongs to a particular family of distributions. It turns out that the negative exponential and the Weibull distributions are the only main distributions that are both accelerated life and proportional hazard models. The difference between the models is that in proportional hazards the applicants most at risk of defaulting at any one time remain the ones most at risk of defaulting at any other time.

Cox (1972) pointed out that in proportional hazards one can estimate the weights **w** without knowing $h_0(t)$, using the ordering of the failure times and the censored times. If t_i , $\mathbf{x_i}$ are the failure (or censored) times and the application variables for each of the items under test, then the conditional probability that customer i defaults at time t_i given R(i) are the customers still operating just before t_i is given by:

$$\exp\{\mathbf{w.x_i}\} h_0(t) \Big/ \sum_{k \in R(i)} \exp\{\mathbf{w.x_k}\} h_0(t) = \exp\{\mathbf{w.x_i}\} \Big/ \sum_{k \in R(i)} \exp\{\mathbf{w.x_k}\} \tag{10}$$

which is independent of h_0. This approach, which does not prejudge the form of the baseline hazard function, is the one that has been most closely explored in the credit context.

One of the disadvantages of the proportional hazards assumption is that the relative ranking among the applicants of the risk (be it of default or early repayment) does not vary over time. This can be overcome by introducing time-dependent characteristics. So suppose $x_1=1$ if the purpose of the loan is refinancing and 0 otherwise. One can introduce a second characteristic $x_2=x_1t$. In one model (Stepanova and Thomas, 1999) with just x_1 involved, the corresponding weight was $w_1=0.157$, so the hazard rate at time t for refinancing loans was $e^{0.157}h_0(t)=1.17h_0(t)$ and for other loans $h_0(t)$. When the analysis was done with both x_1 and x_2, the coefficients of the proportional hazard loans were $w_1=0.32$, $w_2=-0.01$. So for refinancing loans the hazard rate at time t was $e^{0.32-0.01t}h_0(t)$ compared with others $h_0(t)$. Thus in month 1, the hazard from having a refinancing loan was $e^{0.31}=1.36$ times higher than for a non-refinancing loan, while after 36 months, the hazard rate for refinancing was $e^{-0.04} = 0.96$ of the hazard rate for not refinancing. Thus time-by-characteristic interactions in proportional hazard models allow the flexibility that the effect of a characteristic can increase or decrease with the age of the loan.

Survival techniques can also be applied in the behavioural scoring context, though a little more care is needed. Suppose it is u periods since the start of the loan and $\underline{b}(u)$ are the behavioural characteristics in period u , then a proportional hazard model says the hazard rate for defaulting in another t periods time, i.e. t+u since the start of the loan, is $e^{\mathbf{w(u).b(u)}}h_0^u(t)$. At the next period u+1, the comparable hazard rate would be that for t-1 more periods to go, i.e. $e^{\mathbf{w(u+1).b(u+1)}}h_0^{u+1}(t-1)$. Thus the coefficients w(u) have to be estimated separately for each period u, using only the data in the data set that has survived up to period

u. As it stands these coefficients could change significantly from one period to the next. One way of smoothing out these changes would be to make the behavioural score at the last period one of the characteristics for the current period. Another way is to fit a simple curve to explain the time variation in each coefficient $b_i(u)$; so in the linear case one seeks to fit $b_i(u)$ by $a_i + b_i u$. Details of such an analysis can be found in Stepanova and Thomas (Stepanova and Thomas 2001).

7. Conclusions

This paper has reviewed the way consumer risk assessment procedures incorporate the dynamical aspects of consumer behaviour. One can think of application scoring as a way of connecting two snapshots of the consumer together – the first of "his characteristics on applying for a loan" and the second of "his delinquency and default status a year later". In behavioural scoring, the first of these snapshots is replaced by a film clip of the consumer's behaviour over an observation period of six to twelve months but the second snapshot remains. In both application and behavioural scoring this second snapshot seeks to measure the default risk of the consumer twelve months or so after the observation point. Though this risk is time specific, there is a hidden assumption that the relative rankings of default risk hold for some time into the future. However there is no attempt made to measure the default risk of the consumer through the whole of an economic cycle. Given the duration of such cycles and the relative speed with which the characteristics of the borrowing population change, it would not seem possible to do so using the existing methodologies.

The current interest by lenders in developing profit scoring systems means one will need to connect the observation period film clip to an outcome period film clip, since one needs an outcome interval of time to over which to identify the profitability of the customer. Markov chain models are one way of describing the dynamics of the consumer's behaviour in this outcome time period, and are used particularly to estimate delinquency risks either for debt provisioning or for sizing the collections effort. The survival analysis approach on the other hand concentrates on the time dependency of the default risk alone, not on the delinquency states leading up to it. The same approach though can be used to estimate the time dependency of other profit related risks like early repayment or attrition.

It is interesting to compare the similarities in the models used in assessing credit risk in consumer lending and in corporate lending. The credit scoring and behavioural scoring methodologies were used in the 1960s to estimate the likelihood of firms defaulting. Taffler (1982) and Altman (1968) with their ideas of z-scores developed scorecards with accounting ratios as characteristics to measure this risk. They found that they needed different scorecards for different industry sectors and different countries, which meant the population of similar firms was too small for the approach to have the success of credit scoring. Interestingly, the company rating agencies have recently returned to these ideas (Falkenstein et al 2000) to try and get a semi-automatic way of rating all the firms who may want to borrow from financial companies. They are adding subjective estimates of the strength of a firm's management to the accounting ratio characteristics and are experimenting with neural networks and other non-linear classification procedures to try and improve the default risk estimates.

The dynamic Markov chain models described in section five are related to some of the reduced form models introduced by Jarrow and Turnbull (1995) for estimating bond prices. In these models the credit risk that the firm will default on its obligations is modelled using a Markov chain approach based on the credit rating given to the bond by the rating agency. These bond price models also model the interest rate process and the interaction between it and the credit risk. This is in stark contrast to the behavioural and credit-scoring models, which do not even include the current interest rate as a characteristic in their model, let alone model its dynamics. However, as was mentioned, one can introduce the interest rate as a parameter of the transition matrices describing the dynamics of a consumers delinquency status or behavioural score. Again the survival analysis models outlined in the previous section has strong similarities with the proportional hazards approach to credit risk in bonds suggested by Lando (1994). In both cases, one could include the interest rate as one of the characteristics that affects the hazard rate of default.

There are also some examples where the corporate credit risk models and the consumer credit models tackle the same problem but with very different approaches. In the case of mortgage backed securities one can use scoring and survival analysis to build models of the early repayment risk on individual mortgages. Yet these, with their emphasis on the characteristics of the mortgager and the type of property involved, are very different from the models used in corporate finance to price a mortgage backed security, which is nothing but a portfolio of such mortgages. The latter concentrates heavily on modelling the probabilistic nature of the interest rate process and assumes this to be the main driver of early repayment. Similarly there is little intersection between the scoring models used to estimate the default risk in individual consumer loans and the models used to price the risk in portfolios of such loans constructed for securitization reasons. Clearly the "average" of the behavioural scores says something about the expected risk of default in the portfolio, but one needs to get some extra information about the correlation between the risk of defaulting of the separate loans to be able to describe accurately the risk at the portfolio level.

The recent consultative paper from the Basel Committee on Banking Supervision (2001) emphasised the need for banks to have internal models for estimating default risk at the sovereign debt, corporate debt and retail debt levels, and that there be consistency across these internal models. This will undoubtedly lead to a closer connection between the modelling of credit risk at the corporate level and the consumer level in the future, which will be of advantage to both areas.

Acknowledgements

The authors acknowledge the support of an EPSRC visiting fellowship for Prof Scherer. The authors are grateful to the referees for their careful reading and positive suggestions for improving the paper.

References

Altman E.I. (1968), Financial Ratios, Discriminant Analysis and the prediction of corporate bankruptcy, *Journal of Finance* 23, 589-609.

Anderson T.W., Goodman L.A. (1957), Statistical Inference about Markov Chains, *Ann. Math Statist*. 28, 89-109.

Banasik J., Crook J.N. and Thomas L.C. (1999), Not if but when borrowers default, *J. Operational Research Society* 50, 1185-1190.

Bartlett M.S. (1951), The frequency goodness of fit test for probability chains*, Proc. Cambridge Phil. Soc.* 47, 86-95.

Basel Committee on Banking Supervision (2001), *The Internal Ratings Based Approach*, Consultative Document, January 2001.

Bierman H. and Hausman W.H. (1970), The credit granting decision, *Management Science,* 16, 519-532.

Corcoran A.W. (1978), The use of exponentially smoothed transition matrices to improve forecasting of cash flows from accounts receivable, *Management Science,* 24, 732-739.

Cox D.R. (1972), Regression models and life-tables (with discussion), *J. Royal Stat. Soc*., Series B, 74, 187-220.

Cyert R.M., Davidson H.J. and Thompson G.L. (1962), Estimation of allowance for doubtful accounts by Markov chains, *Management Science* 8, 287-303.

Falkenstein E., Carty L., Bard A.K., (2000), RiskCalc for Private companies, Moody's default model, *Moody's Investor Services.*

Fuller D.N. and Scherer W.T. (1999), An exploration and case study of population classification for managed healthcare within a state-based modelling framework, *Working Paper, Dept. Of Industrial Engineering, University of Virginia.*

Frydman H. (1984), Maximum likelihood estimation in the Mover-Stayer model, *J. Am. Stat. Ass*. 79, 632- 638.

Frydman H., Kallberg J.G. and Kao D-L (1985), Testing the adequacy of Markov chains and Mover-Stayer models as representations of credit behaviour, *Operations Research* 33, 1203-1214.

Golabi K., Kalkarni R.B. and Way G.B. (1982), A state wide pavement management system, *Interfaces* 27, 71-88.

Hand D.J. (2001), Modelling consumer credit risk, to appear *IMA J. Management Mathematics* .

Hand D.J., Kelley M.G. (2000), *Lookahead scorecards for new fixed term credit products*, Working paper TR-00-01, Department of Statistics, Imperial College London.

Ho J. (2001), *Modelling Bank customers' behaviour using data warehouses and incorporating economic indicators*, Ph.D. Thesis, University of Edinburgh.

Hoadley B. and Oliver R.M. (1998), Business measures of scorecard benefit, *IMA J. Mathematics Applied in Business and Industry* 9, 55-64.

Hoel P.G. (1954), A test for Markov chains, *Biometrika* 41, 430-433.

Hopper M.A. and Lewis E.M. (1992), Behaviour Scoring and Adaptive Control Systems, in *Credit scoring and Credit Control*, ed. L.C.Thomas, J.N.Crook, D.B.Edelman, Oxford University Press, Oxford, 257-276.

.Jarrow R.A. and Turnbull S. (1995), Pricing derivatives on financial securities subject to credit risk, *J. Finance* 50, 53-86.

Lando D. (1994), *Three essays on contingent claims pricing*, Ph.D. thesis, Cornell University, Ithaca.

Lewis E.M. (1992) *An introduction to credit scoring*, Athena Press, San Rafael, California.

McNab H. and Wynn A. (2000), *Principles and practice of consumer credit risk management*, CIB Publishing, Canterbury.

Narain B. (1992), Survival analysis and the credit granting decision, in *Credit Scoring and Credit Control,* ed. L.C.Thomas, J.N.Crook, D.B.Edelman, , Oxford University Press, Oxford, 109-122

Scherer W.T. and Glagola D.M. (1994), Markovian models for bridge maintenance management, *J. Transportation Engineering* 120, 37-51.

Stepanova M. and Thomas L.C. (1999), Survival analysis methods for personal loan data, Proceedings of Credit Scoring and Credit Control VI, Credit Research Centre, University of Edinburgh, to appear in *Operations Research.*

Stepanova M. and Thomas L.C. (2001), PHAB Scores; Proportional hazards analysis Behavioural Scores, to appear in *J. Operational Research Society*.

Taffler R.J. (1982), Forecasting company failure in the UK using discriminant analysis and financial ratio data, *J. Royal Stat. Soc.*, Series A, 145, 342-358.

Thomas L.C. (1994) Applications and solution algorithms for dynamic programming, *Bulletin of the I.M.A*. 30, 116-122.

Thomas L.C. (2000), A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers, *Int. J. Forecasting* 16, 149-172.

van Kuelen J.A.M., Spronk J. and Corcoran A.W. (1981) Note on the Cyert-Davidson-Thompson Doubtful Accounts Model, *Management Science* 27, 108-112.

Wang K.C.P., Zaniewski J. and Way G. (1994), Probabilistic behaviour of pavement, *J. Transportation Engineering* 120, 358-375

Weiss E.N., Cohen M.A. and Hershey J.C. (1982), An iterative estimation and validation procedure for specification of semi-Markov models with application to hospital patient flow, *Operations Research* 30, 1082-1104.