

Consumer Finance: Challenges for Operational Research

Lyn C Thomas

School of Management

University of Southampton

Southampton

Abstract : Consumer finance has become one of the most important areas of banking both because of the amount of money being lent and the impact of such credit on the global economy and the realisation that the credit crunch of 2008 was partly due to incorrect modelling of the risks in such lending. This paper reviews the development of credit scoring, -the way of assessing risk in consumer finance- and what is meant by a credit score. It then outlines ten challenges for Operational Research to support modelling in consumer finance. Some of these involve developing more robust risk assessment systems while others are to expand the use of such modelling to deal with the current objectives of lenders and the new decisions they have to make in consumer finance.

Introduction

Consumer finance was the sleeping giant of the modern economy until it awoke with a vengeance in 2007 and showed what impact problems with the risk assessment of consumer borrowing and the consequent mis-pricing of financial instruments based on this borrowing could have. Until then despite its importance to the individual consumer, and the fact it was employing an increasing number of those who had trained in Operational Research and statistics, the modelling underlying it was hardly discussed in any finance course and the number of research papers in the area were minute compared with those on the corporate credit market or the pricing of exotic equity based options. This was because the risk models developed in the 1950s and 1960s still seemed to be working well and were surprisingly robust to changes in economic conditions. More emphasis was being put by lenders on the use of Operational Research models in the marketing of these products since the traditional approach of one market “price” (namely the interest rate being charged on the loan) was giving way to variable pricing. At the same time some lenders sought to integrate

all the models into a customer lifetime value framework. These are still challenges for OR in this area but the sub prime mortgage crisis, the failure of the ratings agencies to assess the risk of residential mortgage backed securities, and the consequent credit crunch requires a reassessment of some of the quantitative models which had proved so successful up to then.

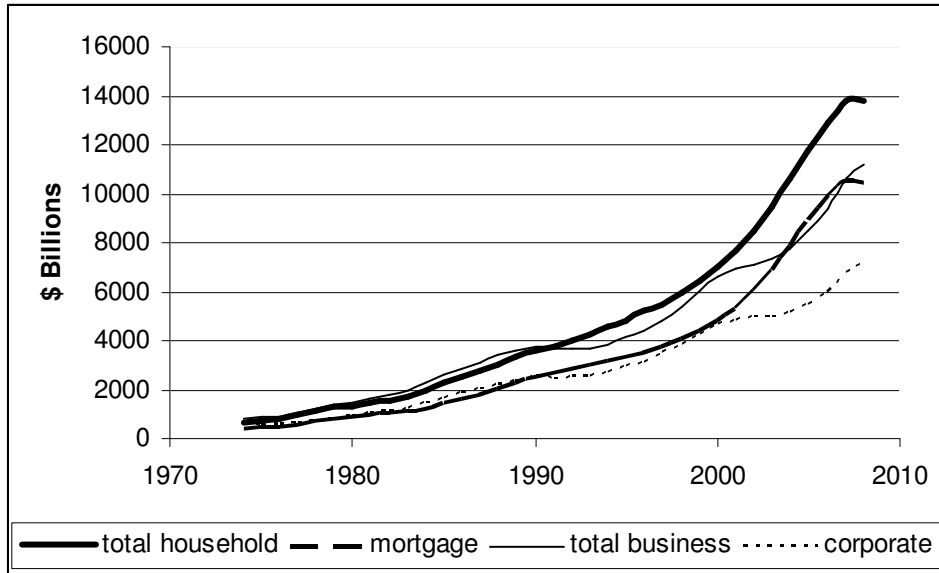


Figure 1: Comparison of US household and business debt

Consumer credit has been around for 4,000 years .There is a Sumerian clay tablet recording how two farmers borrowed money to purchase grain with the promise of paying back more at harvest time. In the Middle Ages the discussion on whether it was right to charge interest on loans not only gave the focal point of a Shakespearean play but exercised both Moslem and Catholic theologians. However it is only in the last fifty years, with the advent of credit cards (first issued in the US in 1958 and then in the UK in 1966) and the growth in home ownership and hence mortgage loans, that consumer credit has become so widespread. Figure 1 show how the total household borrowing in the US overtook that of total business borrowing in the late 1980s and that by 2004 the total borrowing on mortgages had also exceeded the total business borrowing, though that has drawn level again in 2008. Figure 2 similarly shows the growth in consumer borrowing in the UK in the fifteen years from 1992. Borrowing went up more than 350% in that time and even with the housing crisis of 2007 and 2008, the amount outstanding on mortgage loans is still more than £1.2 trillion.

Total UK Personal Debt £bn

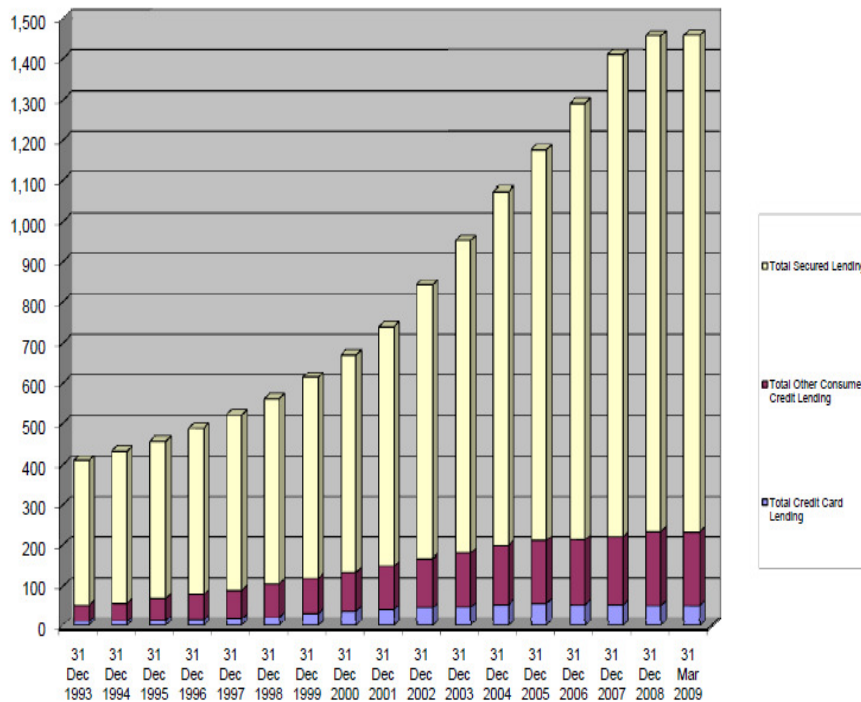


Figure 2: Total consumer borrowing (Calculated by credit action based on Bank of England statistics)

Such growth in consumer lending could not have been possible without an automated approach to assessing the credit risk that the loan to an individual consumer would not be repaid. (In 2007, it was estimated the number of credit cards and debit cards in circulation worldwide exceeded 3 billion. One would need a lot of analysts to subjectively decide whether all those cards should be issued). Moreover laws like the Equal Credit Opportunity Acts in the US have outlawed discrimination in the giving of credit unless there are statistical models which can defend such decisions. These statistically based automated approaches to assessing consumer credit risk go under the name of credit scoring. The models forecast how likely the applicant for credit is to be “Bad” and default on the loan within a given time period. Those borrowers who do not default on the loan within the chosen time period are “Good”. The consumer lending decision can then be modelled as a decision tree. Figure 3 shows a simplified case where the credit score just takes two values- one corresponds to a High chance of being Good (Good Risk), the other to a Low chance of being Good (Bad Risk)

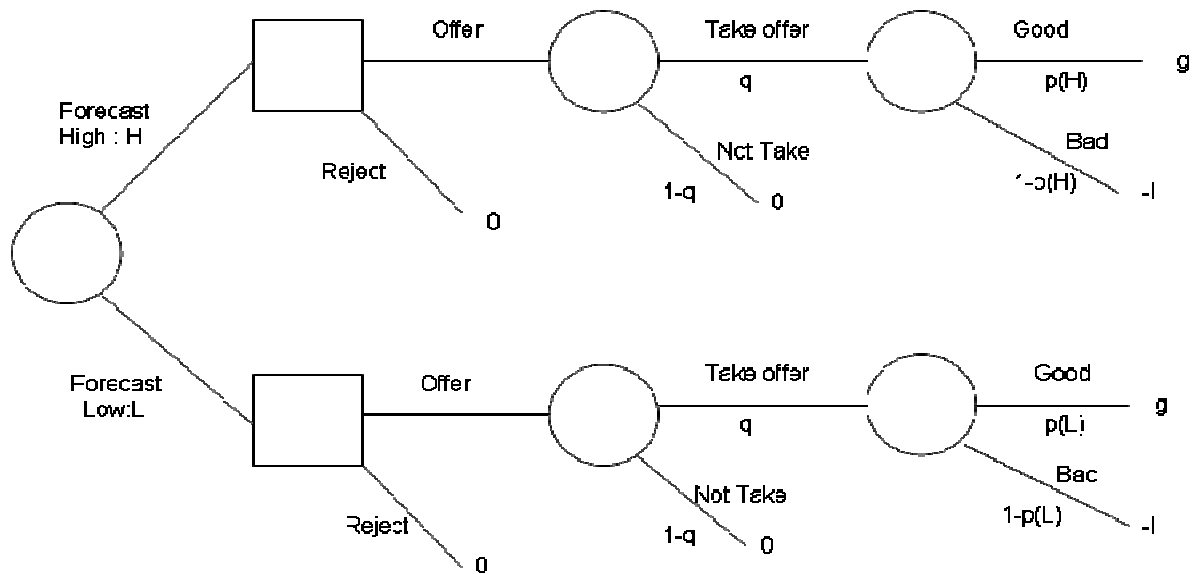


Figure 3: Decision tree of consumer lending decisions

The notation of Figure 3 says that the profit to lender if the loan is repaid is g ; $-l$ is the loss if the loan is not repaid; q is the chance the consumer will take the loan if offered it; $p(H)$ is the probability a consumer rated High will be Good and $p(L)$ is the chance a consumer rated Low will be Good. Then the lender should accept applicants with High credit scores if

$$q(p(H)g + (1-p(H))(-l)) > 0$$

and accept applicants with low credit scores if

$$q(p(L)g - (1-p(L))l) > 0.$$

Credit scoring began in the 1950s when it was realised that statistical classification methods – the first being discriminant analysis (Fisher 1936) – could be used to classify loans into Goods (non defaulting) and Bads (defaulting) using the characteristics of the loan and the borrowers. Initially it was used by mail order companies and finance houses and only after the advent of credit cards did banks start using it – firstly for credit cards, then for personal loans and finally for mortgages. This initial use of credit scoring, which is called application scoring, was to support the decision of whether to grant credit to a new applicant. Its philosophy was pragmatic, in that it only wanted to predict not explain and so used any characteristic that improved the discriminating power of the system. Moreover it concentrated on a very specific risk – the chance a borrower will become 90 days overdue in their repayments in the next 12 months. Whether the loan was profitable to the lender;

whether the borrower would continue to repay beyond this period; how much the borrower used the loan facility; none of these risks were considered. The approach also assumed that the relationship between loan/ borrower characteristics and credit worthiness was stable at least over a four or five year period. It took data on applicants of two years ago, and looked at their performance over the subsequent year. This performance was used to determine whether the applicant was Bad (the specific risk occurred) or Good (it did not occur). This sample was then used to build a classification system which best separated the Goods from the Bads using the characteristics of the loan and the borrower. The standard classification methods result in a scorecard and a cut-off so that those with scores above the cut-off are considered Good (and would be accepted if they apply) and those below it are classified as Bad (and would be rejected if they apply). So a scorecard built on a two year old sample is used to determine which applicants to take for the next few years. After some time, the process is repeated and a new scorecard constructed.

The second variant of credit scoring, behavioural scoring, was introduced in the 1980s when it was felt useful to assess the credit risk of existing customers as well as new applicants. So again the target variable was whether the borrower would default in the next 12 months but now it was possible to use information on the borrower's recent (usually last 12 months) repayment and purchase performance. Such scores are now used by almost all lenders and are routinely updated each month. The most powerful characteristics are whether the borrowers have recently been in arrears and the current information from the credit bureau on their overall credit performance. Although behavioural scoring was an obvious extension of application scoring it was also an opportunity missed. Firstly it is not used to support a specific decision but rather it is used by the lender as part of a customer relationship strategy to determine whether to increase credit limits, seek to up sell or cross sell other products. The aim of these actions though is to improve the profitability of the customer but there might be other measures rather than default risk in the next 12 months which give a better handle on profit. Also behavioural scoring only used static characteristics about the customer's past performance and used these to estimate the customer's status at a fixed time in the future. An alternative would have been to build a dynamic model of how a customer has been performing, which would allow one to forecast the future dynamic behaviour of the customer.

In the past few years, credit scoring had been changing as lenders want credit scoring to support their business objectives of profitability and market share. Lenders want to optimise all the decisions they make about the borrower not just whether or not to offer the borrower a standard loan product. Even in the initial decision, lenders now have a number of variants of a loan product they can offer, be it platinum, gold, silver or standard credit cards, or tracker, fixed rate, and variable rate mortgages, and within each they can decide what credit limit to offer and what interest rate and fee (the price components) to charge. The growth in the internet and the telephone as ways of undertaking the application process means applications are essentially private and so the product can be “customised” to depend on the applicants’ characteristics, allowing for variable pricing. Similarly lenders are more likely to adjust the product or offer alternative or extra products during their relationship with the customer and so are anxious to know what impact such changes will have on the default risk and the profitability of the customer. Lenders want to use “credit scoring” to help make these variable pricing decisions and to determine the long term profitability of a customer under different lender actions. Moreover profitability is as much about marketing as about risk assessment and so there is a need to combine the work done by financial organisations’ marketing and risk assessment OR groups. Currently these groups see themselves as adversaries with one group wanting to take as many applicants as they can and the other to be as discriminating as possible about who they take. The models used by marketers to segment customers and to estimate propensity of purchase are very similar to the ones used by the risk teams to determine how many different scorecards to develop and then to estimate the likelihood of default for each customer.

The other factor which has been affecting credit scoring in the past few years, is the change in banking regulations introduced by the Basel II Accord (BCBS 2005a). Under these new regulations, banks are allowed to use the estimates from their own internal risk rating systems in the formula which determines the minimum capital they have to set aside to cover the credit risk in their lending. Clearly for lending to consumers these internal risk rating systems are application and behavioural scoring systems. In fact it is only worth banks moving to these internal ratings based systems, if they use them for their consumer lending, since the main saving in capital compared with the alternative externally imposed capital ratios is in consumer lending. The Accord requires its ratings to have many of the properties of the existing credit

scoring systems. For example it defines default as 90 days overdue in the next 12 months (though some national regulators such as the Financial Services Authority in the UK had modified this to 180 days overdue). However it also requires much more of credit scoring with its emphasis on validating the probability of default estimates rather than just ensuring the ranking of borrowers is accurate which was how credit scoring systems were previously judged. The Accord also concentrates on the long run probability of default not just the probability of default in the next twelve months; it emphasises the need to stress test the models and it also requires some completely new estimates such as loss given default, which we will return to later.

Defining a credit score

In this paper we outline some of the challenges that these developments in credit scoring are bringing. We also discuss what re-evaluation is needed of the methodology that underpins scoring because of the problems of the last few years in consumer lending and the mis pricing of the securitized products based on such lending. Before doing that it is worth recalling what a credit score is and what properties it has.

We assume that each consumer, be it an applicant in the case of an application score or a current borrower in the case of a behavioural score, can be described by a set of characteristics $\mathbf{x} = (x_1, x_2, \dots, x_m)$, $\mathbf{x} \in \mathcal{X}$ where \mathcal{X} is the set of all possible borrower characteristic combinations. These characteristics include socio-economic data like age and residential status; credit bureau information, like whether the applicant is on the electoral role; and in the case of behavioural scores, performance data like the number of missed payments in the last 12 months. Having decided on what risk is being assessed – say repayments being more than 90 days overdue in next 12 months- those for which that event occurs are Bads and the others are the Goods. A score, $s(\mathbf{x})$, is then a function of the characteristics \mathbf{x} of a potential borrower which can be translated into the probability estimate that the borrower will be Good. The critical assumption in credit scoring is that the score is all that is required for predicting the probability of the applicant being Good. It is like a sufficient statistic. One also usually assumes the score has a monotonic increasing relationship with the

probability of being Good, so if a borrower has a higher score than a second borrower, the first borrower has a higher probability of being Good than the second.

A proper or sufficient score $s(\mathbf{x})$ captures as much information for predicting the probability of a performance outcome, say Good/Bad, as does the original data vector, \mathbf{x} . so that

$$\Pr\{\text{Good} \mid \text{score based on } \mathbf{x}\} = p(G \mid s(\mathbf{x})) = p(G \mid s(\mathbf{x}), \mathbf{x}) = p(G \mid \mathbf{x}) \quad \forall \mathbf{x} \in \mathcal{X} \quad (1)$$

When appropriate we will drop the \mathbf{x} dependence of the score and write

$$p(s) = p(G \mid s(\mathbf{x})) \quad \text{and} \quad 1 - p(s) = 1 - p(G \mid s(\mathbf{x})) = p(B \mid s(\mathbf{x})) \quad \forall \mathbf{x} \in \mathcal{X} \quad (2)$$

One form of a score is the log odds score where

$$s(\mathbf{x}) = \ln \frac{p(G \mid \mathbf{x})}{p(B \mid \mathbf{x})} \quad p(G \mid \mathbf{x}) + p(B \mid \mathbf{x}) = 1 \quad \mathbf{x} \in \mathcal{X} \quad (3)$$

So a log odds score could have values from minus infinity (when $P(G|\mathbf{x})=0$) to plus infinity when ($P(G|\mathbf{x})=1$). Log odds scores are produced when one uses logistic regression to determine the classification scorecard but can be obtained from other approaches by scaling, so it is reasonable to assume a scorecard has such a property.

Specifying the score of an event is equivalent to specifying its probability because we can write the probability in terms of the score:

$$p(G \mid \mathbf{x}) = \frac{e^{s(\mathbf{x})}}{1 + e^{s(\mathbf{x})}} = \frac{1}{1 + e^{-s(\mathbf{x})}} \quad (4)$$

One interesting feature of a log odds score is that it separates out completely the information about the population from the information about the individual borrower being scored. Applying Bayes' rule in the case of the probability of a Good or a Bad having attributes \mathbf{x} with the distribution of Goods and Bads in the population given by p_G and p_B respectively gives

$$P(G \mid \mathbf{x}) = \frac{p(\mathbf{x} \mid G) p_G}{p(\mathbf{x})}; \quad P(B \mid \mathbf{x}) = \frac{p(\mathbf{x} \mid B) p_B}{p(\mathbf{x})} \quad (5)$$

where $p(\mathbf{x})$ is the probability that an applicant will have attributes \mathbf{x} . Applying this in equation (3) gives

$$s(\mathbf{x}) = \ln\left(\frac{p_G p(\mathbf{x}|G)}{p_B p(\mathbf{x}|B)}\right) = \ln\left(\frac{p_G}{p_B}\right) + \ln\left(\frac{p(\mathbf{x}|G)}{p(\mathbf{x}|B)}\right) = \ln o_{pop} + \ln I(\mathbf{x}) = s_{Pop} + s_{Inf}(\mathbf{x}) \quad (6)$$

Thus a log odds score is the sum of a term depending only on the population odds ($s_{pop} = \ln o_{pop}$) and a term which depends on the information on the borrower \mathbf{x} . The first term on the RHS of (6) is the “prior” score -that is the score of a randomly selected individual from the population; this score is then increased or decreased by the score which is based on the data that is unique to a particular individual.

For further details on the basics of credit scoring and the different approaches to building a scorecard one can look at the books by Mays (Mays 1998), McNab and Wynn (McNab and Wynn 2000), Thomas et al (Thomas et al 2002 and 2004), Mays (2004), Anderson (Anderson 2007)and Thomas (Thomas 2009a) and the review papers by Hand and Henley (1997), Thomas (Thomas 2000), Thomas et al (2005) and Crook et al (Crook et al (2007). The type of marketing models that can be used in consumer finance can be found in examples like Lilien and Rangaswamy (2004).

Now it is time to turn to the challenges that Operational Research in Consumer Finance faces.

Challenge 1: Finding the “silver bullet” or is there a better way to build risk assessment systems

Discriminant analysis was the first way that scorecards were built (see Eisenebeis 2004 for a critique of its use in credit scoring) but by the early 1980 , the growth in computer power had meant that logistic regression had taken over as the main way commercial scorecards were built) Anderson 2007, Mays 2004). Other approaches based on linear programming (Freed and Glover 1981, 1986) and maximising divergence (Thomas 2009a) are also used commercially. Another popular alternative is to use classification trees, with its origins both in statistics (Breiman 1984) and machine learning (Quinlan 1993), though of course this ends up not with a scorecard but with groups of customers described by combinations of their characteristics where each group is classified as either Good or Bad . However any classification approach can be applied to the credit scoring problem and so in the past twenty years

researchers have tried neural nets (Desai et al 1997, Malhotra and Malhotra 2002), support vector machines (Huang et al 2007, van Gestel et al 2003, Bellotti and Crook 2009a) , genetic algorithms (Desai et al 1997, Ong et al 2005), nearest neighbour methods (Chatterjee and Barcun (1970), Henley and Hand (1996)) and ant colony optimization (Martens et al 2007). The review paper by Baesens et al (Baesens et al 2009) explains how Operational Research models and data mining methods are used for a number of such classification problems particularly in credit scoring. More sophisticated versions of these regression approaches have been looked at including projection pursuit regression, and multivariate adaptive regression splines (Lee and Chen 2005).

So what methodology gives a scorecard with the best discrimination in credit scoring? What often happens is that the paper that introduces a new method can show that there is some small improvement by using it rather than some existing method, but one is always slightly concerned that this may be down to the expertise of the authors in their own method and the fact they do not take such care with the existing methods. For example many of the newer methods are essentially constructing non-linear scorecards with interactions between the characteristics but experts in the linear approaches to credit scorecard building – logistic and linear regression- tend to know from experience about such interactions and allow for them by building separate scorecards for different segments of the population or by introducing interaction variables. Baesens et al (Baesens 2003b) undertook a careful comparison of different methods and Xiao et al (Xiao et al 2006) compared the more recently applied methods. It is true that some methods performed slightly better than others - neural nets, support vector machines, logistic regression – but the differences were small and often the hypothesis that two scorecards were equally good at discriminating could not be rejected. Moreover in several countries one has to be able to explain why one rejects an applicant for credit and so “black box” methods like neural nets and support vector machines would not be allowed. Thus researchers are looking to see if they can devise classification trees that mimic the performance of the “black box” and so give reasons for assuming the applicant is Bad and should be rejected.(Baesens et al 2003a, Martens et al 2008).

One way of finding an improved risk system is to use a combination of methods. For example there are classification trees where some of the variables are a “score”

obtained using another method. Similarly one might have a regression approach where one characteristic is the different nodes of a classification tree. Another area where researchers are seeking to find improved credit scoring methods - that is find the silver bullet which will be “the” best way of building scorecards - is ensemble methods. This follows Breiman’s introduction of random forests (Breiman 2001) which consists of a large number of classification trees , each built on a subset of the data and only using a subset of the characteristics. A new case is then classified by each of these trees and its predicted class is taken to be that which the majority of the trees predict. This idea of building a large number of models and choosing what the majority predict could be used with all the classification methodologies not just classification trees.

However the idea that a new methodology will produce far better discrimination using existing characteristics than the current methods is questioned by many experts. There is a view (Overstreet et al 1992) that there are a large number of quite different scorecards which have close to the best discrimination possible – the flat maximum effect- and so in the large samples used to build commercial scorecards, it is likely most methods will find one of these almost optimal scorecards. Still that does not stop people trying, though it would be more useful if the experiments were carried out on the sizes of samples – 10,000 to 50,000 - usually used in scorecard building rather than the small samples of less than 1000 which are easily available in the public literature.

Challenge 2: Introducing economics and market conditions into risk assessment systems

The assumption that credit worthiness is time independent over intervals of three or four years meant that credit scores have been built using the socio-demographic characteristics of the borrower, the credit bureau information about the borrower, details of the loan, and even the repayment performance of the borrower on the loan but not anything about the current economic and market conditions. This assumption has been challenged in the last few years first by the Basel Accord which makes a point that its definition of probability of default is a long run average (i.e. averaging over a full economic cycle) and not just the point in time probability of default. This suggests that the probability of default does vary as economic conditions vary even if

the credit worthiness of the borrower is not changing and has required some ingenuity by lenders to translate a credit score which is clearly a point in time (PIT) estimate into the Through the cycle (TTC) estimate. Secondly the detailed investigations of the sub prime mortgage crisis showed (Demyank and Van Hemert 2008) that the credit scores changed as the economic conditions worsened. This is only to be expected if we recall the decomposition of the credit score in (6). If we include the time at which the score is being used there, then what we require at time t is the score

$$s(\mathbf{x}, t) = s_{Pop}(t) + s_{Inf}(\mathbf{x}, t) \quad \text{What we have is } s(\mathbf{x}, t_0) = s_{Pop}(t_0) + s_{Inf}(\mathbf{x}, t_0) \quad \text{where } t_0 \text{ is the}$$

time at which the sample on which the scorecard was built was performing. One could possibly argue that $s_{Inf}(\mathbf{x})$ is independent of t , even though that is highly unlikely but there is no way $s_{Pop}(t)$ cannot depend on the current economic and market conditions. There are some recent suggestions of how to include these economic conditions, either directly into a regression scorecard (Zandi 1998), or using survival analysis (Malik and Thomas 2008, Bellotti and Crook 2009b). Similar ideas (Tang et al 2007) were previously used to identify how the likelihood of purchasing financial products depends both on the characteristics of the customer and on the economic conditions. The use of interaction terms and time-dependent coefficients which proved so successful there can obviously be taken across to building economy based credit scorecards.

Challenge 3: Dealing with new ways of assessing what is a Good customer

The traditional way of defining a Bad was a borrower who became 90 days overdue in the next 12 months. When personal bankruptcy rules became easier in certain countries, notably the US, it became apparent that the performance of borrowers before they sought bankruptcy was different from those who just defaulted on their loan. Thus bankruptcy scores were developed where a Bad was someone who went bankrupt in the next 12 months.

Although a system which assessed the profitability of the customer is the aim of many lenders, this is proving hard to implement. Instead what has happened is that lenders score separately a number of the events that affect profitability. Attrition scores assess whether the borrower will cancel the loan product shortly. Usage scores assess how much a borrower will use the loan product. Propensity scores assess how likely it

is the lender can up sell or cross sell other products to the borrower. Li and Hand (Li and Hand 2002) suggested that instead of assessing default risk directly, one should try to predict future values of other aspects of the borrower's performance, like the balance on the account, and then from these estimate the risk of default. Such an indirect approach requires both that the intermediate elements can be predicted well and that there is a strong relationship between them and the default risk. It does have the advantage though that it may be possible to use these intermediate components to estimate profitability as well as default risk.

One real change in defining Good/Bad in the last few years, is the use of survival analysis ideas to allow the estimation of a borrower's default risk over any future time horizon not just a fixed 12 months. In survival analysis one is interested in estimating the default hazard rate, $h(t)$, where $h(t)\delta t$ is the conditional probability of default in $(t, t+\delta t]$ given there has been no default in $(0, t]$. So if T is the time when default occurs, $P_B(t)$ is the probability that there has been default by time t , ($P'_B(t)$ its derivative) and $P_G(t)=1-P_B(t)$, then:

$$h(t) = \Pr\{t \leq T < t + \delta t \mid T \geq t\} \triangleq \frac{P'_B(t)}{1 - P_B(t)} = -\frac{d}{dt} \ln(1 - P_B(t)) = -\frac{d}{dt} \ln(P_G(t)) \quad (7)$$

This is not the probability that a borrower will default at a time t into the loan but rather the probability that given the borrower is still active at time t he will default in the next period of time. It is easy then to see that given the hazard function we can calculate the probability of default over any time period because

$$H(t) = \int_0^t h(u) du \quad P_G(t) = 1 - P_B(t) = e^{-H(t)} \quad (8)$$

If one uses the proportional hazards or accelerated life models of survival analysis one is then able to obtain a score which describes the "risk" of a consumer defaulting over any and all time horizons. In the proportional hazard model, the hazard function for default at time period t into the loan for a borrower with characteristics \mathbf{x} decomposes into the product of the baseline hazard function times an enhanced risk due to the borrower's characteristics, namely

$$h(t, \mathbf{x}) = e^{\mathbf{w} \cdot \mathbf{x}} h_0(t) = e^{-s(\mathbf{x})} h_0(t) \quad (9)$$

So $s(\mathbf{x})$ can be considered as a risk score in that the higher the score the less likely the borrower is to default. This model can work both as a parametric model where the

baseline hazard function is of a specific family of distributions or semi parametrically using the results of Cox (Cox 1972). Cox showed one can first calculate the score without making any assumptions about the distribution and then use the Kaplan Meier approach to estimate the empirical distribution for $h_0(t)$ that best fits the data.

In the accelerated life model one can only use the parametric approach but the assumption is that the probability of a borrower with characteristics \mathbf{x} not defaulting before time t ($P_G(t, \mathbf{x})$) is given by

$$\begin{aligned} P_G(t, \mathbf{x}) &= P_{G,0}(e^{\mathbf{w} \cdot \mathbf{x}} t) = P_{G,0}(e^{-s(\mathbf{x})} t) \\ \text{or } h(t) &= e^{\mathbf{w} \cdot \mathbf{x}} h_0(e^{\mathbf{w} \cdot \mathbf{x}} t) = e^{-s(\mathbf{x})} h_0(e^{-s(\mathbf{x})} t) \end{aligned} \quad (10)$$

where again $s(\mathbf{x})$ is the equivalent of a risk score.

These ideas have been developing over the past decade (Banasik et al 1999), Stepanova and Thomas, 2001, 2002), and are now being taken on board by practitioners. Survival analysis has also been used to build scorecards when only a few months data is available (Hand and Kelly 2001). As mentioned in Challenge 2 survival analysis can also be used to introduce economic conditions into scorecards.

One advantage of the survival analysis approach is that the competing risk idea means one can use the same data to estimate several different events. In the competing risk approach one has several ways in which a loan could finish- default, early repayment, normal repayment - and one can model each of these separately using the fact that as far as a default is concerned a borrower who pays off early at time t has a history censored at that time. This competing risk approach can be expanded in two directions. One can model purchasing as well as attrition and default events separately and then seek to combine them to get a customer lifetime value approach (Challenge 10). Alternatively one can concentrate on default only but recognise that default can occur for different reasons – financial naivety, loss of employment, fraud, marital breakdown for example – and seek to model the time until default for these different reasons separately before finally combining them using the competing risk idea.

Thus there seems to be a great deal more research that is required to develop more appropriate Good/Bad assessments both in terms of expanding from default to

profitability and in removing any pre defined time horizon on the time over which the customer is assessed.

Challenge 4: Variable and Risk based pricing

One way that lenders are seeking to increase their profit is to offer generic loan products such as credit cards but to tailor the details of the product for each individual. In credit cards, this would mean varying the interest rate charged, the credit limit offered, whether an annual fee is charged and whether bonuses like air miles are given for purchase made with the card. This is possible because the use of the internet and the telephone as application channels mean the application process is much more private and so varying offers can be made without applicants being aware of what is being offered to others.

One of the simplest schemes would be to adjust the interest rate charged r , to be a function of the probability, p , of the applicant being a Good. For a log odds score equation (4) shows how this probability is related to the credit score of the applicant. Consider a very simple example where 1 unit is lent, the cost of capital for the lender is r_F (the risk free rate), the loss given default (the fraction of the amount outstanding at default which is finally lost) is l_D , and the lender will charge an interest rate $r(p)$ which is related to the probability p of the applicant being a Good. If the take probability or response rate of an applicant to a loan offer with interest rate r is $q(r)$ then the expected profit to the lender of making an offer r is

$$\text{Max}_r E[P(r)] = (q(r)((r - r_F)p - (l_D + r_F)(1 - p))) \quad (11)$$

Differentiating (11) and setting the derivative to zero gives

$$\begin{aligned} q'(r)((r - r_F)p - (l_D + r_F)(1 - p)) + q(r)p &= 0 \\ \Rightarrow r(p) &= r_F - \frac{q(r)}{q'(r)} + \frac{(l_D + r_F)(1 - p)}{p} \\ \Rightarrow r(s) &= r_F - \frac{q(r)}{q'(r)} + (l_D + r_F)e^{-s} \end{aligned} \quad (12)$$

where we assume that s is a log odds application score so $s = \ln\left(\frac{p}{1-p}\right)$.

The reality is that the take probability q is a function of r and p . This is because of adverse selection (Ausubel 1999, Calem et al 2006) in which more Bads apply for

consumer credit at higher interest rates than might be expected. There are also affordability issues since the interest rate charged can affect the ability of the borrower to repay as was seen in the sub prime mortgage crisis. There many of the borrowers only defaulted when the interest rates went from the initial low rates to the higher rates that came in after two or three years of the loan. There is very little mathematical modelling of what are appropriate variable rate functions to charge apart from Phillips's book (Phillips 2005). Similarly there needs to be much more empirical work on what are appropriate take probability functions. In particular how does the take probability vary according to the risk score of the applicant and the rate charged by the lender. Since there are so many combinations that could be considered there are experimental design problems for any lender in order to obtain this sort of information efficiently. Other factors must also be important such as the rates being charged by other lenders, whether the product offers other features, such as air miles or free travel insurance and in the case of revolving credit, whether the applicants believe they will be transactors (pay off their balance every month) or revolvers (and so have balances on which interest is charged).

The problem of finding the optimal price to sell a product at has been around for many years. There are two main approaches. One is to estimate the response function (the take probability) as above while the second is to model the situation as a game. Such games could involve a number of buyers – the borrowers in this case – and sellers – the lenders, and the use of game theory to model such pricing situations has a long history from Edgeworth's work on market games in 1881 (Edgeworth 1881) to Gibbens and Kelly's work on pricing the internet (Gibbens and Kelly 1999)

Challenge 5: Expanding approaches to deal with new forms of credit granting

As well as new modelling challenges in existing forms of credit granting there are new types of loans that need different risk assessment systems to those that have worked for personal loans, credit cards and mortgages. The two that are attracting most interest at present are micro-credit and pay day loans.

Microcredit involves giving very small loans to those in poverty in order to help them develop a business which will sustain them and their family and so bring them out of poverty. It began in the Indian subcontinent but is now being used by many other countries and is even recognised by major international banks as a significant source

of future lending. The United Nations declared 2005 to be the International Year of MicroCredit. Clearly standard risk assessment systems can not work for people who have no history of being advanced credit previously and no involvement with a banking system. Yet there is a need to ensure that the credit loaned is repaid, even if the time periods involved may be very long – several years if not decades- and there is a need to assess both the character of the individual and the potential of the idea which the loan will initially fund. Recently there has been some initial work on how one would need to modify standard credit scoring systems to deal with these questions (Mok 2008).

At the other extreme of time scale is pay day loans. Payday loans are small, very short-term loans with extremely high interest rates that are effectively advances on a borrower's next pay packet. The loan is taken out usually at the middle or towards the end of the month and the lender is given a post dated cheque or a way of accessing the borrower's current account on the day the pay cheque is paid in at the end of the month. This is a much faster moving environment than that for normal loans, since the loans are of such short durations, and their repayment depends on the borrower's ability and desire to pay back the loan that month. So proven ability to handle such short term loans, and the local economic situation are important features. Thus one needs to build scorecards that can respond very quickly to changes in economic and market behaviour and to immediate changes in borrower behaviour and circumstances. Moreover such loans are increasingly receiving special legislation which requires proof that their risk assessment systems are robust.

Challenge 6: Meeting the regulatory challenge , particularly that in the Basel Accord

As mentioned earlier the introduction of the new banking regulations- the Basel II Accord (BCBS 2005a)-, concerning the amount of capital banks need to set aside to cover their risk, has had a major impact on credit scoring. Introduced in Europe in 2007/8 , the US in 2009 and in scheduled for most countries between 2008 and 2012, it was a response to the distortions in lending caused by the first Accord of 1988 rather than a response to the credit crunch. Although it should have has some affect on the lending that precipitated the sub prime mortgage crisis if it had been in effect then, it would not have dealt with the liquidity risk or the fact that some lenders thought

securitization meant they can absolve themselves of the risks of their lending. It is likely that some governments will now impose tighter regulations than those proposed in the Accord. However the idea that banks need to build models of the credit risk of their lending and the output of these is used to set their capital requirements– the internal based rating approach – will remain.

The Accord is presenting four challenges to credit scoring – the internal ratings approach to consumer lending. The first is the need to validate the probability of default predictions that the scorecard makes rather than the relative ranking of the borrowers, which was what is important in deciding which applicants for credit to accept. . So one needs to be confident in the translation of score to probability of default and to use the standard chi square and normal distribution type tests to validate the model by backtesting to compare actual numbers of defaults with predicted ones. (BIS Working paper 14, 2005b) Since there is clearly some dependence between defaults of different individuals, and often the number of defaults are very low, one needs to develop sophisticated models to cope with these problems (Benjamin et al 2006).

A second challenge is the Accord requires estimates of the long run average of the 12 month default rate (the TTC default rate) for a segment of borrowers while a credit score estimates the default rate in the next 12 months (the PIT estimate). Translating from one to the other highlights the time dependency of a score which we outlined in Challenge 2. If $s(t, \mathbf{x})$ is a log odds score at time t for a borrower with characteristics \mathbf{x} then the probability of defaulting in the next 12 months $p_t(B, \mathbf{x})$ starting at t is

$$p_t(B, \mathbf{x}) = \frac{1}{1 + e^{s(t, \mathbf{x})}} \quad (13)$$

This is the PIT estimate but what one needs to do is get a TTC estimate which if the cycle is of length T starting say at time t_0 would be

$$\bar{p}(B, \mathbf{x}) = \int_{t_0}^{t_0+T} \frac{dt}{1 + e^{s(t, \mathbf{x})}} \quad (14)$$

This involves estimating how the score $s(t, \mathbf{x})$ changes over time, which brings us back to Challenge 2: It also presupposes that the score to probability of default transformation stays as a log odds transformation and ignores what happens when scores are recalibrated during the cycle.

A third problem is the Basel Accord's instance on stress testing which means predicting the future performance of a portfolio of loans under extreme economic conditions. One can do this by sensitivity analysis where one changes the value of one of the factors that impacts on the model or by scenario analysis. In the latter approach one identifies a combination of the overall conditions which can lead to poor economic performance. Although there have been several surveys of what stress testing banks currently do (BIS 2005, FSA 2005), these point out to the lack of a consistent stress testing methodology for credit risk as opposed to market risk. The critical issue is how to build a model of the credit risk of portfolios of consumer loans which includes economic and market conditions and so can then be run under the extreme scenarios suggested by the regulators. This is so important we identify it as a separate challenge (Challenge 7) and discuss it further there. Researchers are beginning to address different ways of building models of the credit risk for portfolios of consumer loans which can then be used for stress testing (Breeden 2007, Breeden et al 2008b, Rosch and Schuele 2008, Malik and Thomas 2009)

Similarly the fourth issue that the Basel Accord has highlighted, the need to model the recovery rate RR (or alternatively the loss given default LGD, where $RR=1-LGD$) of what percentage of a defaulted loan will subsequently be recovered is also so important that it deserves to be considered as a separate challenge (Challenge 8).

Challenge 7: Modelling the credit risk of portfolios of consumer loans

Credit scoring has proved very successful at assessing the relative risk of individual borrowers defaulting. The previous discussion though highlighted the credit rating agencies failure to assess the risk of consumer asset backed securities and the Basel requirements to stress test portfolios of consumer loans. Both these show the need for such risk assessment also to be modelled at the portfolio level. Portfolio level credit risk models were developed more than a decade ago for corporate loans with models that allowed the correlation in share prices to be surrogates for the correlation in defaults. This is not possible nor sensible for portfolios of consumer loans as default there does not depend on the value of assets but on cash flow considerations and personal attitudes to debt. However that does not prevent building credit risk models for portfolios of consumer loans which have strong parallels with the corporate

portfolio models (Thomas 2009b). As was suggested in Challenge 6, several models are being developed all of which include economic conditions as part of the model. By applying Monte Carlo simulation using different future economic scenarios one can then use such model to estimate portfolio level default rates. The types of models developed so far include reputation based models (Andrade and Thomas 2007), dual time dynamics (Breeden 2007, Breeden and Thomas 2008a), survival analysis (Bellotti and Crook 2009b, Malik and Thomas 2008,), and correlation models with added economic variables (Rosch and Scheule 2003). Given the amount of research that has gone into corporate credit risk models, one suspects that there will be considerable more research into these consumer equivalents given the realisation by bankers now of how much more is being lent to households than to companies.

Challenge 8: Modelling Loss Given Default and the Collection Process

There had been little analytic modeling of the collections process for any form of lending until the advent of the Basel Accord. The Accord though requires banks to estimate LGD (Loss Given Default) (Bennett et al 2005) for all loan segments whether they have yet defaulted or not. LGD is related to the recovery rate, RR (i.e., the percentage of the debt outstanding which the collections department recovers), by $LGD = 1 - RR$. Before this there had been some work on estimating recovery rates in corporate lending since these affect the price of risky bonds. The edited book by Altman et al (Altman et al 2002) outlines the mainly regression-based models that seek to relate recovery rates to economic factors and characteristics of the loan and the defaulter in the corporate setting. The work on modeling the collections process for mortgage lending (Lucas 2006) is directly motivated by Basel. This model splits the problem into whether the mortgaged property needs to be repossessed and then into forecasting what price the property will be sold for. Such two stage models could also be used for other secured loans like car finance .

For unsecured consumer credit, Matuszyk et al (Matuszyk et al 2007) have recognized that the recovery rate depends both on decisions by the lender as well as the uncertainty about the borrower's ability and intention to repay. They used a decision tree approach to model the strategic level decision of whether to collect the debt in house, use an agent or sell off the debt. Modelling the amount recovered overall (or under one of these strategies) in terms of the characteristics of the debtor and the loan

is proving to be very difficult. The initial approaches have looked at linear and logistic regression, non linear transformation so as to fit Beta or log log distributions, mixture models (especially to identify the “won’t pay” (LGD =1), and even quantile regression ideas (Somers and Whittaker 2007). All seem to give correlations between actual and predicted values of no better than 0.1 to 0.2. Moreover, the data that banks are now storing systematically on the outcomes of their collections process is being used to develop models of the sequence and timing of the collections operations so as to optimize the recovery rate (De Almeida Filho et al 2008). So not only is it proving difficult to get reasonable estimates of LGD and RR using existing data, but building models to optimize or at least improve the collections process is likely to mean that recovery rates in the future will be significantly improved on those found in this data. Thus currently LGD modeling is like estimating a moving target

Challenge 9: Developing combined marketing and risk assessment models that help with the operations management of borrower’s accounts

One of the most surprising aspects of consumer lending in most financial organisations is the lack of integration between the marketing and credit risk groups. Both are interested in maximising the profit for the organisation by making decisions about potential and actual customers; both use statistical methods to segment the population and to predict how likely the customer is to perform certain events- be it purchasing a new financial product or defaulting on an existing loan. The methods used are very similar- almost all the methods mentioned in challenge 1 could be applied to build marketing prediction models. Both groups use the same data about a customer to build their models and yet rarely are combined models built.

The book by Beck and Siegel (Beck et al 2001) outlines the way marketing is used in consumer lending but there are surprisingly few integrated models which include risk and marketing features in consumer lending. In fact Burez and van den Poel (2008) produce a churn model in a paper entitled resolving the conflict between the sales and credit department. One could argue the pricing models of challenge 4 are a start but the marketing aspects of the model are not widely used , apart from the work on multiple features in credit cards (Thomas et al 2006). Buckinx et al (Buckinx et al 2007) use the transactional information to estimate the customer loyalty to the organisation, while van den Poel and Lariviere (2004) model which product features

prevent customers churning to another organisation. There seem to be so many obvious benefits in seeking to integrate the ideas and the models in the two areas. For example there are some marketing models that seek to assess the “emotions” of the customer from their interactions with the company (Coussement et al 2009) but there is no risk assessment models which includes the customers “emotions”.

Challenge 10; Developing valid customer lifetime value models when lifetime means lifetime.

This final challenge is an obvious extension of Challenge 9. The overall goal of marketing and credit risk modelling is to improve the profitability of the customer to the financial organisation by improving customer relationship management. To do this one needs to estimate customer lifetime value. Whereas in many retail environments the horizon may be just until the next purchase or possibly just for a few years, in the consumer finance area lifetime can really mean lifetime – pension products for example. Thus one needs to build lifetime value models that can cope with the changes in economic and market conditions over long time intervals as well as forecasting the changes in the customer’s situation and priorities. Tang et al (Tang et al 2007) built a survival analysis model which included the interactions between economic and socio demographic variables to estimate changes in the purchases of pension products. Donkers et al (2007) made a comparison of a number of different types of customer lifetime value models using insurance industry data, while in Verhoef and Donkers (2001) they made the comparison between choice based probit models and potential value regression type models. Baesens et al (Baesens et al 2004) used Bayesian network classifiers to estimate the parameters of where in the life cycle a customer might currently be. Benoit and van den Poel (2009) have used quantile regression to estimate customer lifetime value. All these models concentrate on the purchase aspects – time to and value of next purchase and churn- and do not include the default risk elements which can affect profitability in a major way. With approaches such as the competing risk idea in survival analysis it should be possible to combine these two major factors which affect customer profitability.

Conclusion

Given the turmoil in the financial markets during 2007 and 2008, which has at last made practitioners and researchers realise how larger a proportion of the banking industry is based on consumer lending, there is no question that research in this area will be very active for the foreseeable future. In particular, the lack of models for the credit risk of portfolios of consumer loans and not modelling how economic conditions affect credit scores is now recognised as having exacerbated the credit crunch of 2008/9. Since many researchers have for more than a decade addressed these problems in corporate lending, it is reasonable to expect they will expand their research to the consumer lending case.

The tremendous increase in computer storage capacity and the requirement of the Basel Accord that banks have sufficient historical data to validate their credit scoring models have meant that banks are now willing and able to store much more consumer finance data over much longer periods than they used to do. This will prove a vital tool in meeting several of the challenges outlined previously. For example, up to five years ago, most banks had hardly any data on the outcome of their collections and recoveries process, but the need to estimate Loss Given Default for all consumer loans means that such data is now carefully recorded and analysed.

Having seen what impact the failure to control the risks in consumer lending have had on the world economy, regulators and bankers will want to develop suitable models (and have enough analysts to build and monitor them) to control these risks in the future – or at least for the next decade. This should mean that consumer finance will have a much higher profile in university Finance and Operational Research courses in the future, so that entrants to the finance industry are aware of the needs and the challenges of building models to solve the problems in this area. This article has sought to identify some of these challenges.

References

Altman E., Resti A., Sironi A., (2005), *Recovery Risk*, Risk Books, London

Anderson R, (2007), *The Credit Scoring Toolkit Theory and Practice for Retail Credit Risk Management and Decision Automation*, Oxford University Press, Oxford.

- Andrade F.W.Muniz de, Thomas L.C. (2007), Structural models in consumer credit, *European Journal of Operational Research* 183, 1569-1581.
- Ausubel L.M., (1999) Adverse selection in the credit card market , Working Paper, University of Maryland
- Baesens B. Mues C. Martens D. Vanthienen J. (2008). 50 years of datamining and OR: upcoming trends and challenges. *Journal of the Operational Research Society*. 60, S16-S23.
- Baesens, B., Setiono, R., Mues, C., Vanthienen, J. (2003a). Using Neural Network Rule Extraction and Decision Tables for Credit-Risk Evaluation. *Management Science*, 49(3), 312-329.
- Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J. and J. Vanthienen (2003b). Benchmarking state-of-the-art classification algorithms for credit scoring. *Journal of the Operational Research Society*, 54, 6, 627-635.
- Baesens B., Verstraeten G., van den Poel D, Egmont-Petersen M., van Kenhove P., Vanthienen J., (2004), Bayesian Network Classifiers for identifying the slope of the customer lifecycle of long-life customers, *European Journal of Operational Research* 156, 508-523.
- Banasik J., Crook J.N., Thomas L.C. (1999) Not if but when borrowers default, *J. Operational Research Society* 50, 1185-1190.
- Bank of International Settlements, (2005), Stress testing at major financial institutions: survey results and practice, CGFS Publication 24, Basel
- Basel Committee on Banking Supervision (2005a, comprehensive version 2006), International convergence of capital measurement and capital standards – a revised framework, Bank for International Settlements, Basel
- Basel Committee on Banking Supervision, (2005b), Studies on the validation of Internal rating systems, Working Paper 14, Basel.
- Beck R.E., Siegel S.M., (2001), *Consumer Lending*, Kogan Page
- Bellotti T., Crook J.N, (2009a), Support vector machines for credit scoring and discovery of significant features , *Expert Systems with Applications* 36,3302-3308
- Bellotti T., Crook J.N. (2009b). Credit scoring with macroeconomic variables using survival analysis. *Journal of the Operational Research Society*, to appear
- Benjamin N, Cathcart A, Ryan K. (2006), Low default portfolios: a proposal for conservative estimation of default probabilities, Financial Services Authority , London.
- Bennett R.L., Catarineu E., Moral G., (2005), Loss Given Default validation, Studies on the validation of Internal rating systems, pp60-76, Working Paper 14, Basel Committee on Banking Supervision, Basel
- Benoit D.F., van den Poel D, (2009), Benefits of quantile regression for the analysis of customer lifetime value in a contractual setting: an application in financial services, to appear in *Expert Systems with Applications*

- Breeden J.L., (2007), Modeling data with multiple time dimensions, *Computational Statistics and Data Analysis*, 51, 4761-4785.
- Breeden, J. L., Thomas, L. C., (2008a) The relationship between default and economic cycle for retail portfolios across countries. *Journal of Risk Model Validation* 2 (3), 11-47.
- Breeden, J. W., Thomas, L. C., McDonald, (2008b). Stress testing retail loan portfolios with dual-time dynamics. *Journal of Risk Model Validation* 2 (2) , 43-62.
- Breiman L., Friedman J.H., Olshen R.A., Stone C.J. (1984) Classification and regression trees, Wadsworth, Belmont, California.
- Breiman L., (2001), Random Forests, *Machine Learning* 45, 5-32.
- Buckinx W., Verstraeten G., van den Poel D., (2007), Predicting Customer loyalty using the Internal transactional Database, *Expert Systems with Applications* 32, 125-134
- Burez J, Van den Poel D, (2008), Separating financial from commercial churn: a modeling step towards resolving the conflict between the sales and credit department, *Expert Systems with Applications* 35, 497-514
- Calem P. S., Gordy M. B. and Mester, L. J., (2006). Switching costs and adverse selection in the market for credit cards: New evidence, *Journal of Banking & Finance*, 30, 1653-1685.
- Chatterjee S., Barcun S., (1970), A nonparametric approach to credit screening, *J. American Statistical association* 65,150-154.
- Coussement K., van den Poel D, (2009), Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers, *Expert Systems with Applications* 36, 6127-6134.
- Cox D.R., (1972), Regression models and life tables (with discussion), *Journal of Royal Statistical Society, Series B*, 187-220
- Crook J.N., Edelman D.B, Thomas L.C.,(2007). Recent Developments in Consumer Credit Risk assessment. *European J. Operational Research*, 183, 1447-1465
- De Almeida Filho A.T., Mues C., Thomas L.C. (2008),Optimizing the collections process in consumer credit , Working Paper Centre for Risk Research University of Southampton.
- Demyanyk Y. and Van Hemert O. (2008), Understanding the Subprime Mortgage Crisis, Working paper (2008), available at the Social Science Research Network. See ssrn.com/abstract=1020396.

Desai V.S, Conway D.G., Crook J.N., Overstreet G.A. (1997), Credit scoring models in the credit union environment using neural networks and genetic algorithms, IMA Journal of Mathematics Applied in Business & Industry 8, 323-346

Donkers B, Verhoef P, Jong M, (2007), Modeling CLV: A test of competing models in the insurance industry , Quantitative Marketing and Economics 5, 163-190

Edgeworth, F. Y. (1881) "Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences" Kegan Paul and Co., pp. viii., 150

Financial Services Authority (2005), Stress Testing, Discussion Paper 05/02, London

Fisher R.A., (1936) The use of multiple measurements in taxonomic problems, Annals of Eugenics 7, 179-188.

Freed N., Glover F., (1981) A linear programming approach to the discriminant problem, Decision Sciences 12, 68-74.

Freed N., Glover F., (1986), Evaluating alternative linear programming models to solve the two-group discriminant problem, Decision Sciences 17, 151-162.

Gibbens R.J. and Kelly F.P. (1999), Resource pricing and the evolution of congestion control, Automatica, 35, 1969-1985,

Hand D.J. and Henley W.E. (1997) Statistical classification methods in consumer credit scoring: a review. Journal of the Royal Statistical Society, Series A, **160**, 523-541.

Hand D.J. and Kelly M.G. (2001) Lookahead scorecards for new fixed term credit products. Journal of the Operational Research Society, **52**, 989-996.

Henley W.E. and Hand D.J., (1997) Construction of a k-nearest neighbour credit scoring system. IMA Journal of Mathematics Applied in Business and Industry, **8**, 305-321.

Huang C-L, Chen M-C, Wang C-J, (2007), Credit scoring with a data mining approach based on support vector machines, Expert Systems with Applications, **33**, 847-856..

Lee T-S and Chen I-F, (2005), A two-stage hybrid credit scoring model using artificial neural networks and multivariate adaptive regression splines. Expert Systems with Applications. **28**, 743-752.

Li H.G. and Hand D.J. (2002) Direct versus indirect credit scoring classifications. Journal of the Operational Research Society, **53**, 1-8.

Lilien G.R., Rangaswamy A., Marketing Engineering, Trafford, Victoria (2004)

Lucas A, (2006), Basel II Problem Solving,
<http://www3.imperial.ac.uk/portal/pls/portal/active/docs/1/7287866.PDF>

- McNab H and Wynn A (2000). Principles and Practice of Consumer Credit Risk Management. CIB Publishing, Canterbury.
- Malhotra R., and Malhotra D.K., , (2002), Differentiating between good credits and bad credits using neuro-fuzzy systems. *European Journal of Operational Research*. v136 i1. 190-211.
- Malik M., Thomas L.C. ,(2008), Modelling credit risk of portfolio of consumer loans, to appear in *Journal of Operational Research Society* , Working Paper, Centre for Risk Research, School of Management, University of Southampton, Southampton.
- Malik M., Thomas L.C., (2009), Transition matrix models for consumer credit ratings, Working Paper , CORMSIS, University of Southampton
- Martens, D., De Backer, M., Haesen, R., Vanthienen, J., Snoeck, M. and Baesens, B. (2007). Classification with Ant Colony Optimization. *IEEE Transactions on Evolutionary Computation*, Vol. 11, No. 5, 651-665.
- Martens D., Huysmans J, Setiono R, Vanthienen J., Baesens B, (2008) Rule Extraction from Support Vector Machines: An Overview of Issues and Application in Credit Scoring. *Rule Extraction from Support Vector Machines*, 33-63
- Matuszyk A., Mues C., Thomas L.C. (2007), Modelling LGD for unsecured personal loans; Decision Tree approach, Working Paper CORMSIS 07-07, School of Management, University of Southampton , to appear in *Journal of the Operational Research Society*.
- Mays E, (1998), *Credit Risk Modeling, Design and Application*, Fitzroy Dearborn publishers, Chicago.
- Mays E., (2004), *Credit Scoring for Risk Managers, The Handbook for Lenders*, Thomson South Western, Mason, Ohio
- Mok J-K, (2008) *Process Scoring for Micro Loans*
<http://www.few.vu.nl/stagebureau/stage/stageverslagen/stageverslag-mokg.pdf>
- Ong C.S., Haung J.J., Tzeng G., (2005), Building credit scoring models using genetic programming, *Expert Systems with Applications* 30, 507-518.
- Overstreet G.A., Bradley E.L., Kemp R.S. Jr, (1992), The Flat Maximum Effect and Generic Linear Scoring Models: A Test, *IMA Journal of Management Mathematics* 4, 97-109.
- Phillips R.L., (2005), *Pricing and Revenue Optimization*, Stanford Business Books, Stanford, California.
- Quinlan J.R., (1993) *C4.5: Programs for Machine Learning*, Morgan Kaufman, San Mateo, California

Rosch D., Scheule H.(2003) Forecasting retail portfolio credit risk, *Journal of Risk finance* 5, 16-32

Rosch D., Scheule H., (2008), *Stress Testing in Financial Institutions*, Risk Books, London

Somers M, Whittaker J (2007), Quantile regression for modelling distributions of profit and loss. *European Journal of Operational Research* 183, 1477-1487

Stepanova M., Thomas L.C., (2002) Survival analysis methods for personal loan data, *Operations Research* 50, 277-289,

Stepanova M., Thomas L.C., (2001), PHAB scores: Proportional hazards analysis behavioural scores, *J. Operational Research Society* 52, 1007-1016.

Tang, L.L., Thomas, L.C., Thomas, S. and J-F. Bozzetto (2007). It's the Economy Stupid: Comparison of Proportional Hazards Models with Economic and Socio-demographic variables for estimating the purchase of financial products. *International Journal of Bank Marketing*, 25 , 22-38.

Thomas, L.C. (2000). A survey of credit and behavioural scoring; Forecasting financial risk of lending to consumers. *International Journal of Forecasting*, 16, 149 - 172.

Thomas L.C., (2009a), *Consumer Credit Models, Pricing, Profit and Portfolios*, Oxford University Press, Oxford.

Thomas L.C., (2009b) Modelling the Credit Risk for Portfolios of Consumer Loans: Analogies with corporate loan models ,*Mathematics and Computers in Simulation* 20, 2525-2534.

Thomas L.C., Edelman D.B., Crook J.N., (2002) *Credit Scoring and its Applications*, SIAM, Philadelphia, US.

Thomas L.C., Edelman D.B., Crook, J.N., (2004), *Readings in Credit Scoring*, Oxford University Press, Oxford.

Thomas, L.C., Oliver, R.W. ,Hand D.J. (2005). A survey of the issues in consumer credit modelling research. *Journal of the Operational Research Society*, 56, 1006-1015.

Thomas L.C., Jung K.M., Thomas, S.D.A., Wu Y., (2006),*Modelling Consumer Acceptance Probabilities*, *Expert Systems and their Applications* 30, 507-518,

Van den Poel D., Lariviere B., (2004), Customer Attrition Analysis for financial services using proportional hazard models, *European Journal of Operational Research* 157, 196-217

Van Gestel T., Baesens B., Suykkens J.A.K, Van den Poel D.,Baestaens D.E, Willikens Mmet al., (2006), Bayesian kernel based classification for financial distress detection. European Journal of Operational Research, 172. 979-1003.

Verhoef, P.C. and Donkers, A.C.D., (2001). Predicting Customer Potential Value: an application in the insurance industry, Research Paper ERS-2001-01-MKT Revision_, Erasmus Research Institute of Management (ERIM)

Xiao, W., Zhao, Q., Fei, Q., (2006). A comparative study of data mining methods in consumer loans credit scoring management. Journal of Systems Science and Systems Engineering 15 (4), 419-435.

Zandi M., (1998), Incorporating Economic Information into Credit Risk Underwriting in (1998), Credit Risk Modeling, Design and Application, ed E.Mays, Fitzroy Dearborn publishers, Chicago, pp155-168..