# Credit price optimisation within retail banking

SE Terblanche*  T de la Rey†

## Abstract

The willingness of a customer to pay for a product or service is mathematically captured by a price elasticity model. The model relates the responsiveness of customers to a change in the quoted price. In addition to overall price sensitivity, adverse selection could be observed whereby certain customer segments react differently towards price changes. In this paper the problem of determining optimal prices to quote prospective customers in credit retail is addressed such that the interest income to the lender will be maximised while taking price sensitivity and adverse selection into account. For this purpose a response model is suggested that overcomes non-concavity and unrealistic asymptotic behaviour which allows for a linearisation approach of the non-linear price optimisation problem. A two-stage linear stochastic programming formulation is suggested for the optimisation of prices while taking uncertainty in future price sensitivity into account. Empirical results are based on real data from a financial institution.

## 1  Introduction

In recent years there has been a significant shift by some industries to move away from cost based pricing, where the price of a product or service is based on the cost plus some fixed profit margin, to a more flexible demand-based pricing strategy [18]. *Demand-based pricing* is done by taking into account the willingness of a customer to pay for a product or service, *i.e. price elasticity*. The responsiveness of the quantity demanded of a product or service to a change in its price is a measure of elasticity and it is commonly referred to as a *price response function*. Examples of typical price response functions are the linear, the constant-elasticity and the s-shaped price response function [13].

Knowing more about the customer would likely improve the predictive power of a response function. Cross and Dixit [5] state that the key to *customer-centric pricing* is to set prices that accurately reflect the perceived value of products per customer segment, where

---

*Corresponding author: Centre for Business Mathematics and Informatics, North-West University, Potchefstroom, email: fanie.terblanche@nwu.ac.za

†Centre for Business Mathematics and Informatics, North-West University, Potchefstroom

a customer segment is a grouping of customers having similar characteristics and product preferences. Agarwal and Furgeson [1] presents bid-response models for customised business-to-business bid pricing and show that by making use of customer segmentation an increase in profits can be expected. Phillips [15] suggests a pricing approach where price levels are determined per customer segment while taking price sensitivity into account.

In retail banking, specifically consumer credit, the pricing approach followed for many years was limited to *risk-based pricing* [7]. The dependence on risk-based pricing could be attributed to the uncertainty in the expected revenue and costs associated with consumer credit. Caufield [4] argues that risk-based pricing is the lending industry's version of cost-based pricing and that an increase in profit of between 10 to 25 percent could be expected with a profit-based pricing approach. Such an approach would typically combine risk-based pricing and demand-based pricing in an attempt to maximise profits.

The problem being addressed in this study is the pricing of consumer credit products in retail banking. In addition to general price sensitivity, *adverse selection* is an important characteristic in retail credit that is likely to have a significant impact on pricing [19]. In literature various definitions of adverse selection can be found. For instance, there is a distinction made between adverse selection on observable information and adverse selection on hidden information [2]. Phillips and Raffard [14] make the same distinction but refer to direct and indirect adverse selection. In this paper adverse selection on observable information (direct adverse selection) is applicable and an example of this is when low risk customers are more sensitive to an increase in price compared to high risk customers [14]. Therefore, according to Thomas [20], adverse selection needs to be taken into account as part of risk-based pricing since it influences the interaction between the quality of the customers and the probability of them taking up credit products.

The literature contains empirical evidence of the existence of price elasticity and adverse selection in retail credit. More specifically, Park [12] finds that for the credit card industry a decrease in demand is associated with an increase in price. This is also the case for the credit industries in less developed economies [10]. Ausubel [2] also finds clear evidence of adverse selection within the credit card industry, and in Einav *et al.* [6], an empirical model of demand for subprime credit is developed that takes adverse selection into account. Applying their model on detailed cost data they find that optimal prices dictate lowering down payment requirements for low risk customers and increasing it for high risk customers.

Most of the literature on the topic of pricing, as outlined above, focusses on determining the factors that influence price setting and the relationships that may exist between consumer behaviour and pricing. Only recently have there been efforts to formalise the retail credit price optimisation problem and the challenges faced with obtaining optimal solutions. Phillips [16] considers the pricing problem for credit consumers that entails determining optimal prices per pricing segment according to an objective function that combines the net interest income with price sensitivity. The log-concavity property of the said objective function allows for the efficient generation of optimal prices provided that any additional side constraints preserve convexity of the feasible region. No numerical results are presented in this paper. Oliver and Oliver [11] provide a numerical algorithm to find optimal prices to maximise return on equity by considering price response and

default risk. Their approach is based on the solution of non-linear differential equations.

The price optimisation model considered in this paper is based on the work by Phillips [16] and takes uncertainty in future price sensitivity into account. To the best of the authors' knowledge uncertainty in future price sensitivity has not been considered previously in any study concerned with retail credit price optimisation. Furthermore, a response model is suggested that overcomes non-concavity and unrealistic asymptotic behaviour. This allows for a linearisation of the retail price optimisation problem making it more tractable to obtain optimal solutions using standard linear programming technology. In addition, the linearisation allows for a stochastic linear programming approach for solving the credit price optimisation problem with uncertainty in future price sensitivity. The empirical results presented in this paper are based on real data from the South African retail bank Absa, a subsidiary of Barclays Bank Plc.

In the next section empirical evidence is provided that support the use of a stochastic programming framework. In §3, the income function used to approximate the net present interest income is introduced and details are provided of the proposed response function. A linear approximation of the proposed response function is provided and incorporated into a linear programming model for solving the credit price optimisation problem with multiple customer segments. In §4, the basic problem is extended to cater for uncertainty in future price sensitivity by formulating the credit price optimisation problem as a linear stochastic programming problem. Empirical results are provided in §5 that highlights the benefits of following a stochastic programming approach based on real data. Finally, summary remarks and a conclusion are provided in §6.

## 2 Price sensitivity and adverse selection

In this paper the optimisation problem of solving the retail credit price optimisation problem is addressed by taking uncertainty in future price sensitivity into account. The output of this model is thus the optimal prices that will be quoted to prospective customers, *i.e.* the loan interest rates, such that the expected interest income to the lender, discounted with the effect of price elasticity and adverse selection, will be maximised. In this section, evidence of the existence of price elasticity and adverse selection will be provided and a case will be made for using a stochastic programming framework to solve the retail credit price optimisation problem.

To illustrate the effect of price sensitivity and the effect of adverse selection, empirical tests were performed using data obtained from a financial institution in South Africa over a period of three years. The variables contained within the data set included:

- whether a customer took up a loan ($Y = 1$) or not ($Y = 0$)
- the quoted interest rate ($r$)
- the repurchase rate[1] ($r_0$)
- the loan amount ($a$)
- the loan term ($n$)

---

[1]The interest rate at which commercial banks can borrow money from the Reserve Bank.

- the probability of default $(p)$
- the loan application date.

The probability of a customer taking up a loan is expressed as the following response function obtained from fitting a logistic regression model

$$R(r, a, n, p) = 1 / \left( 1 + e^{(-(\beta_0 + \beta_1 r + \beta_2 a + \beta_3 n + \beta_4 p + \beta_5 r_0))} \right), \tag{1}$$

with $\beta_1$ to $\beta_5$ the regression coefficients that are estimated through the maximum likelihood method. Note that for ease of illustration, interaction terms between the different variables have been omitted in the formulas. They were, however, included during the model building process. Furthermore, instead of following the customary approach of modelling the margin $r - r_0$, the repurchase rate $r_0$ is considered separately and is shown in the results to feature in some of the interaction terms. The response function $R(r, a, n, p)$ gives the probability that a customer with a probability of default[2] of $p$ will take up a loan of size $a$, over a term of $n$ and with a quoted price of $r$, provided that the current repurchase rate is $r_0$. In subsequent sections, the notation $R(r, a, n, p)$ will be used to obtain a two dimensional response function in terms of the variable $r$ by supplying constant values for $a$, $n$ and $p$ to the logistic regression model (1).

A stepwise logistic regression was performed (p-value of 5%) and a $c$-statistic of 0.608 was obtained. Although a perfect model would have yielded a $c$-statistic of one (0.5 for a random model) it should be noted that the data set under consideration is limited in the number of variables and with additional variables capturing information such as demographics, application turn-around time, macro economic factors, *etc.* an improved c-statistic may be possible. It should be noted that in this paper caution is taken in drawing conclusions about the underlying model other than the inverse relationship between take-up and price, which has already been established in literature. An improved $c$-statistic would clearly provide more confidence in the optimal prices obtained. However, it is not a requirement in order to prove the benefit of the proposed optimization framework, which is evident from the results.
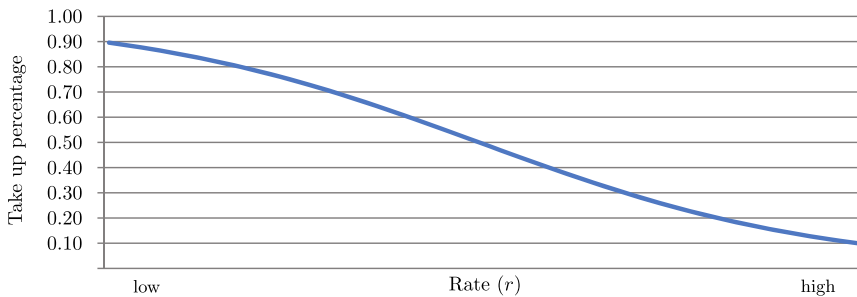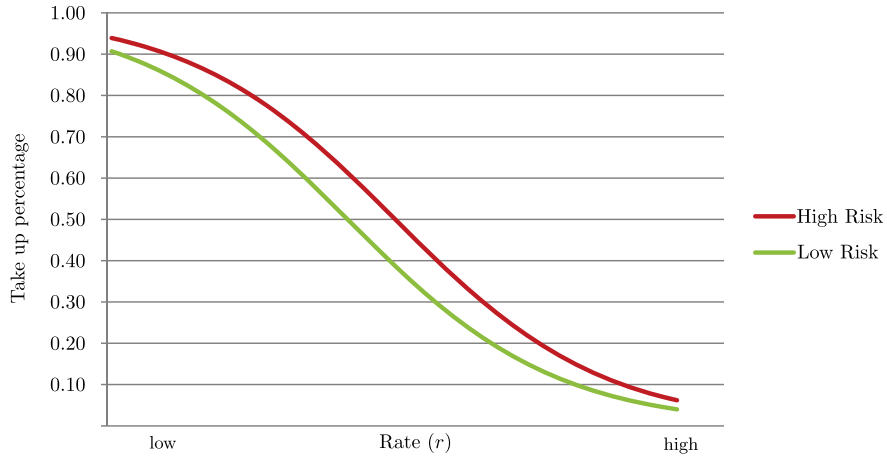


**Figure 1:** *A response graph of price versus take-up of a loan. The units of r are omitted due to a confidentiality agreement.*

In Figure 1, the response graph of price versus take-up is shown, generated from the response function $R(r, a, n, p)$ by substituting averages from the data set for $a$, $n$ and $p$.

---

[2]A financial term describing the likelihood that a customer will be unable to meet their debt obligation.

From the graph it is clear that price elasticity exists since lower take-up rates are associated with an increase in price. To illustrate the effect of adverse selection, different levels of probability of default ($p$) were substituted into the response function $R(r, a, n, p)$ for low risk and high risk customers respectively. As may be seen from Figure 2, the take-up for high risk customers was higher compared to low risk customers for the same quoted price, implying that adverse selection does indeed exist for this data set.



**Figure 2:** *A response graph of price versus take-up for different risk categories.*

Another aspect to consider is the economic conditions over time. A graph is provided in Figure 3 of how the repurchase rate in South Africa has changed over time. It is anticipated that with changing economic conditions there might be a change in sensitivity towards credit prices. In order to test this hypothesis in a very pragmatic way two points in time have been identified that partitions the data set into three periods. The first period coincides with the the first five months during which the repurchase rate dropped significantly, as can be seen from Figure 3. The next period coincides with the subsequent 18 months during which a moderate rate of decrease in repurchase rate can be observed. For the last 13 months the repurchase rate remained unchanged. The logistic regression model (1) was fitted using the data partitions for each of the three time periods. The average repurchase rates from each period were used for $r_0$ in each of the fitted models. For the first time period a c-statistic of 0.621 was obtained, for the second time period a c-statistic of 0.618 was obtained and for the last time period a c-statistic of 0.6 was obtained.

The response functions fitted to each of the time periods are illustrated in Figure 4. The lack of sensitivity towards price increases in the first period could be attributed to an optimistic economic outlook due to the steep decline in repurchase rate within a short period of time. Irrespective of the reason for this phenomenon, it is clear that economic conditions could have an effect on price elasticity. It is, therefore, a reasonable assumption to take potential future price sensitivity into account when determining prices to be quoted to prospective customers. For instance, to determine prices at this point in time one could consider the three response graphs depicted in Figure 4 as potential future scenarios with
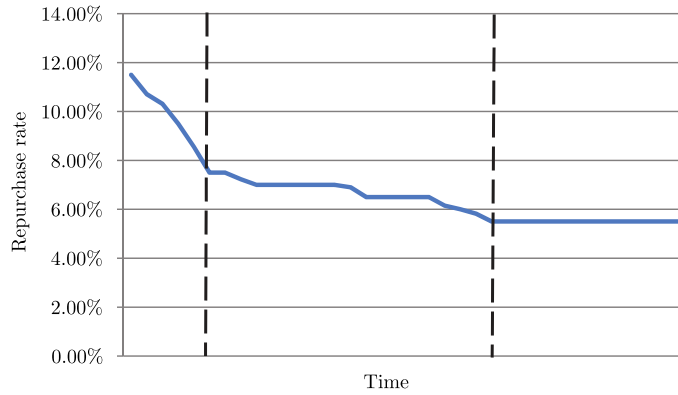
**Figure 3:**  *A graph of the repurchase rate over time. The dashed lines indicate a possible change in economic conditions.*

respect to price sensitivity. That is, in the near future the current repurchase rate level may either drop significantly, or only moderately or remain unchanged as was historically the case. Note that the proposed optimisation approach below will allow us to also incorporate potential future scenarios that are not captured as part of the historic data. For instance, another potential future scenario to consider might be that the current repurchase rate will increase in the future, especially if it is currently at a low level.
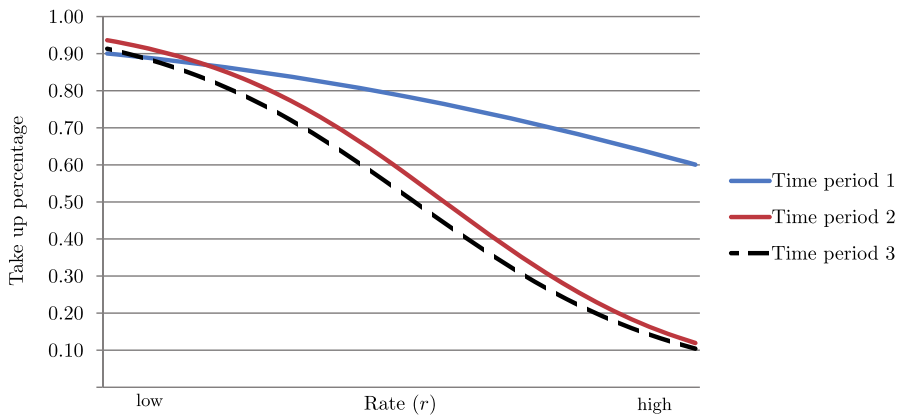


**Figure 4:**  *A response graph of price versus take-up for different time periods.*

In view of the above, the optimisation question at hand is, therefore, to determine optimal price levels by considering the uncertainty in future price elasticity which could be represented by a finite set of potential future scenarios. Preceding the details of the proposed model that addresses this problem, the fundamental building blocks of the credit price optimisation problem is set out below.

# 3   A concave linear approximation of the objective function

Consider a simple loan with a fixed annual rate. Let the probability of default for a customer be denoted by $p$ and the loss given default by $\delta$. Furthermore, let $a$ be the loan amount, $n$ the term (in months) and $r$ the price (annual interest rate). By denoting $r_0$ as the annual repurchase rate, *i.e.* the cost of funding the loan, an approximation of the net present interest income [16] is given by,

$$I(r, a, n, p) = na(r/12 - r_0/12) - ap\delta. \tag{2}$$

The profitability of the customer is expressed in terms of the approximated income which is $na(r/12 - r_0/12)$ minus a cost of risk of $ap\delta$. The attractiveness of the approximation $I(r, a, n, p)$ is that the function is firstly, linear in the rate $r$ and secondly, instead of relying on a sequence of probability of defaults over time, it is written in terms of an overall probability of default $p$ [16]. For the remainder of this paper we assume that $\delta = 1$.

The approximate net present income function (2) can now be generalised to accommodate a customer segmentation approach. In practice prices are determined per customer segment in order to differentiate prices according to product and customer characteristics. For instance, an obvious segmentation scheme for credit is to let customers with similar credit scores applying for loans having similar terms and loan amounts be in the same segment. Let $\mathcal{C} = \{1, 2, \ldots, |\mathcal{C}|\}$ be the index set of all customer segments. By denoting $p_c$, $a_c$ and $n_c$ as the mean probability of default, the mean loan size and the mean term for a customer segment $c \in \mathcal{C}$, the approximate net present income for the segment as a function of the mean rate $r_c$ is

$$\tilde{I}(r_c, a_c, n_c, p_c, v_c) = v_c n_c a_c (r_c/12 - r_0/12) - v_c a_c p_c \delta, \tag{3}$$

with $v_c$ the number of loan applications (volume) for customer segment $c \in \mathcal{C}$. The assumptions underlying (3) is that a 100% take-up is expected by all the customers in the segment for the quoted price $r_c$. In order to reflect the fact that income is conditional on customer take-up and to adjust the income function accordingly, we turn our attention to price elasticity.

In contrast to Phillips [16], this paper does not suggest fitting the model (1) for each customer segment $c \in \mathcal{C}$. The motivation is that data availability in some of the customer segments may lead to response functions with poor predictive power. The alternative is to obtain a single response function $R(r, a, n, p)$ that is fitted by taking the segment averages, $a_c$, $n_c$ and $p_c$, for each customer segment $c \in \mathcal{C}$ as input and by modelling the target variable as $Y_c = \sum_{i \in \mathcal{I}(c)} Y_i / v_c$ with $\mathcal{I}(c)$ the case indices belonging to customer segment $c \in \mathcal{C}$. That is, the resulting data set will have $|\mathcal{C}|$ number of cases. It is anticipated that the average take-up for a customer segment $c \in \mathcal{C}$ given by $R(r_c, a_c, n_c, p_c)$ will have a better smoothing effect over segments with limited cases. Considering the income function (3) and the response function (1), the resulting credit price optimisation problem that maximises the approximate net interest income per customer segment is defined as

$$\max_{r_c \geq 0} \sum_{c \in \mathcal{C}} I(r_c, a_c, n_c, p_c, v_c) R(r_c, a_c, n_c, p_c). \tag{4}$$
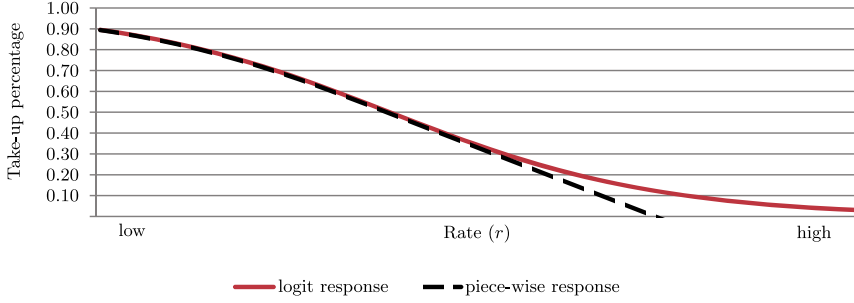
**Figure 5:** *The effect of fitting the piece-wise response function $\tilde{R}(r, a, n, p)$.*

The optimisation problem (4) is an unconstrained problem and solving it using standard non-linear optimisation methods will produce a unique solution since the term $I(r, a, n, p)$ is linear and $R(r, a, n, p)$ is an increasing failure rate distribution [15]. Incorporating constraints into the problem would still result in optimal solutions provided that the solution space remains a convex set [3]. In a practical setup one of the most useful constraints to consider in credit price optimisation is the volume constraint

$$v_c R(r_c, a_c, n_c, p_c) \leq V, \tag{5}$$

with $V$ an upper limit on the proportion of customers defined by the segment $c \in \mathcal{C}$. This could be used to limit the volume of customers having a specific risk profile. It should be noted, however, that by adding a constraint of the form (5) to the optimisation problem (4), the solution space would be a non-convex set since the function $R(r, a, n, p)$ is defined to be neither convex nor concave. Apart from this the response function $R(r, a, n, p)$ also has an unrealistic infinite support with respect to the price variable $r$. In an attempt to address both these issues the response function

$$\tilde{R}(r, a, n, p) = \begin{cases} 1/(1 + e^{-\phi(r)+\alpha}) & \text{if } \phi(r) \geq \alpha, \\ 0.5 + 0.25\phi(r) - 0.25\alpha & \text{if } \alpha - 2 \leq \phi(r) < \alpha, \end{cases} \tag{6}$$

is suggested, with $\phi(r) = \beta_0 + \beta_1 r + \beta_2 a + \beta_3 n + \beta_4 p + \beta_5 r_0$ and $\alpha$ is a shifting parameter. Note that potential interaction terms in $\phi(r)$ have been omitted for ease of illustration. The response function $\tilde{R}(r, a, n, p)$ has the desired property that it is concave with respect to the price variable $r$ on the domain $\phi(r) \geq \alpha - 2$ and it intersects zero due to the linear function $0.5 + 0.25\phi(r) - 0.25\alpha$ that is tangent to the logistic function $1/(1 + e^{-\phi(r)+\alpha})$ in its inflection point. Figure 5 shows the effect of fitting the model $\tilde{R}(r, a, n, p)$ to the data. Statistical results (not provided here) showed that there is a marginal improvement in using this new adjusted response function compared to using the ordinary logit function in terms of goodness of fit.

It should be noted that, although $\sum_{c \in \mathcal{C}} \tilde{I}(r_c, a_c, n_c, p_c, v_c)\tilde{R}(r_c, a_c, n_c, p_c)$ is a concave objective function, obtaining an optimal solution poses a problem to most existing non-linear solvers due to the domain dependent definition of $\tilde{R}(r, a, n, p)$. In order to resolve this issue and to make it implementable for standard convex optimisation technology, a linearisation approach is followed whereby the new concave response function $\tilde{R}(r_c, a_c, n_c, p_c)$, for

a customer segment $c \in \mathcal{C}$, is approximated with piece-wise linear functions with respect to $r_c$. Let the support $0 \leq r_c \leq 1$ be divided into intervals indexed by $\mathcal{K} = \{1, 2, \ldots, 1/w\}$ with $0 < w \leq 1$ the interval width. For each interval $k \in \mathcal{K}$ the response function is approximated with a linear function having a slope of $m_{ck}^R$. In addition, the product $I(r_c, a_c, n_c, p_c, v_c)\tilde{R}(r_c, a_c, n_c, p_c)$ constituting the objective function for a customer segment $c \in \mathcal{C}$, is approximated with linear functions with the slopes of these functions denoted by $m_{ck}^I$ for $k \in \mathcal{K}$. Introducing the incremental price variables $0 \leq r_{ck} \leq w$ and the take-up variables $t_c$ for each of the customer segments $c \in \mathcal{C}$, the linear price optimisation problem (LPOP) is obtained with the objective to

$$\text{maximise} \quad \sum_{c \in \mathcal{C}} \sum_{k \in \mathcal{K}} m_{ck}^I r_{ck} \tag{7}$$

$$\text{subject to} \quad t_c - \sum_{k \in \mathcal{K}} m_{ck}^R r_{ck} = \tilde{R}(0, a_c, n_c, p_c) \qquad c \in \mathcal{C}, \tag{8}$$

$$t_c \geq 0 \qquad c \in \mathcal{C}, \tag{9}$$

$$0 \leq r_{ck} \leq w \qquad c \in \mathcal{C}, k \in \mathcal{K}. \tag{10}$$

The objective function (7) maximises an approximation of $\tilde{I}(r_c, a_c, n_c, p_c, v_c)\tilde{R}(r_c, a_c, n_c, p_c)$ for each customer segment $c \in \mathcal{C}$. Note that the constant intercepts of the net interest income approximations, $\tilde{I}(0, a_c, n_c, v_c, p_c, v_c)\tilde{R}(0, a_c, n_c, p_c)$, do not have to feature in the objective function. Furthermore, the constraint set (8) is only used to assign values to the auxiliary take-up variables $t_c$. Although the take-up variables $t_c$ do not feature in the objective function they can now be used in other linear side constraints since their values depend on the optimal prices $r_c^*$. Specifically, it is shown in the section below how the take-up variables $t_c$ are used in the formulation of the volume constraints (5).

The above formulation of LPOP is a linear representation of a concave maximisation problem and solving it will produce an optimal solution $r_k^*$ for each $k \in \mathcal{K}$ such that

$$\tilde{R}(r_c^*, a_c, n_c, p_c) \approx \tilde{R}(0, a_c, n_c, p_c) + \sum_{k \in \mathcal{K}} m_{ck}^R r_{ck}^*$$

and

$$\tilde{I}(r_c^*, a_c, n_c, p_c, v_c)\tilde{R}(r_c^*, a_c, n_c, p_c) \approx \tilde{I}(0, a_c, n_c, v_c, p_c, v_c)\tilde{R}(0, a_c, n_c, p_c) + \sum_{k \in \mathcal{K}} m_{ck}^I r_{ck}^*,$$

with $r_c^* = \sum_{k \in \mathcal{K}} r_{ck}^*$ for each customer segment $c \in \mathcal{C}$.

## 4   A stochastic programming framework

To determine optimal price levels by considering the uncertainty in future price elasticity a two-stage stochastic programming approach is followed. The basic mechanism of this framework is to adjust the price to be quoted in such a way as to optimise net interest income expressed in terms of potential future realisations of response (see Higle [9] for a

very practical approach in explaining the two-stage mechanism). To achieve this a scenario based approach is followed whereby for each scenario $s \in \mathcal{S}$, with $\mathcal{S} = \{1, 2, \ldots, |\mathcal{S}|\}$ — the index set of all scenarios, the response function

$$\tilde{R}_s(r, a, n, p) = \begin{cases} 1/(1 + e^{-\phi(r,s)+\alpha}) & \text{if } \phi(r, s) \geq \alpha, \\ 0.5 + 0.25\phi(r, s) - 0.25\alpha & \text{if } \alpha - 2 \leq \phi(r, s) < \alpha \end{cases} \tag{11}$$

is fitted, with $\phi(r, s) = \beta_{0s} + \beta_{1s}r + \beta_{2s}a + \beta_{3s}n + \beta_{4s}p + \beta_{5s}r_0$ (interaction terms not shown). The implication of adding a scenario index to the regression coefficients above is that uncertainty within future take-up rates is represented by a finite number of response functions $\tilde{R}_s(r, a, n, p)$, with $s \in \mathcal{S}$. Obtaining different sets of betas for each of the scenarios $s \in \mathcal{S}$ can be accomplished in various ways. For instance, by using the partitioning in the above section that is based on the change in repurchase rate over time, a response function representative of the response behavior of customers within each time period can be estimated. In addition, hypothetical scenarios could be created to simulate potential future scenarios that may not be captured in existing data.

Keeping with the above notation, let $r_c \geq 0$ be the decision variable denoting the optimal price to quote customer segment $c \in \mathcal{C}$. This variable is, within the context of stochastic programming, referred to as the first stage decision variable since the prices to be used in future quotes are determined up front. The scenario dependent take-up variables $t_{cs}$ are referred to as the second stage variables since their values depend on the possible future behaviour of customers as defined by the scenario dependent response functions $\tilde{R}_s(r, a, n, p)$, with $s \in \mathcal{S}$.

The formulation of the credit price optimisation problem can now be reformulated as a stochastic programming problem by considering the linearisations

$$t_{cs} = \tilde{R}_s(r_c, a_c, n_c, p_c) \approx \tilde{R}_s(0, a_c, n_c, p_c) + \sum_{k \in \mathcal{K}} m^R_{cks} r_{ck}$$

for each customer segment $c \in \mathcal{C}$ and for each scenario $s \in \mathcal{S}$, with $m^R_{cks}$ the slopes of the linear functions approximating the response function and

$$\tilde{I}(r_c, a_c, n_c, p_c, v_c)\tilde{R}_s(r_c, a_c, n_c, p_c) \approx \tilde{I}(0, a_c, n_c, v_c, p_c, v_c)\tilde{R}_s(0, a_c, n_c, p_c) + \sum_{k \in \mathcal{K}} m^I_{cks} r_{ck},$$

with $m^I_{cks}$ the slopes of the linear functions approximating the net interest income in the objective function. For both linearisations $r_c = \sum_{k \in \mathcal{K}} r_{ck}$. In order to state the complete credit price optimisation problem as a linear stochastic programming problem and to incorporate practical side constraints, some additional notation is required.

A risk grading $g \in \mathcal{G} = \{1, 2, \ldots, |\mathcal{G}|\}$ is a classification according to the average probability of default, allowing the definition of the index set $\mathcal{C}(g)$ of all customer segments having a risk grading $g$. Conversely, the mapping $\mathcal{G}(c)$ can be used to retrieve the risk grading for a customer segment $c \in \mathcal{C}$. The ordering $g_1 < g_2$ with $g_1 = \mathcal{G}(c_1)$ and $g_2 = \mathcal{G}(c_2)$ implies the ordering of the probabilities of default $p_{c_1} < p_{c_2}$ with $c_1, c_2 \in \mathcal{C}$. Furthermore, a similarity index $h \in \mathcal{H} = \{1, 2, \ldots, |\mathcal{H}|\}$ is assigned to each customer segment such that the index set $\mathcal{C}(g, h)$ denote all the segments that have a risk grade of $g \in \mathcal{G}$ and the

same loan amount and term, *i.e.* having the same similarity index $h \in \mathcal{H}$. The mapping $\mathcal{H}(c)$ can be used to retrieve the similarity index of a customer segment $c \in \mathcal{C}$. For the purpose of imposing constraints on the retail credit price optimisation problem, the following additional parameters are required:

1. Let $L_g$ denote the lower bound (in percentage) on the take-up volume of loans over all customer segments classified as having a risk grading $g \in \mathcal{G}$.
2. Let $U_g$ denote the upper bound (in percentage) on the take-up volume of loans over all customer segments classified as having a risk grading $g \in \mathcal{G}$.

Let $\rho_s$ denote the probability of response scenario $s \in \mathcal{S}$ realising in future. The resulting stochastic linear price optimisation problem (SLPOP) is formulated as

$$\text{maximise} \quad \sum_{c \in \mathcal{C}} \sum_{s \in \mathcal{S}} \sum_{k \in \mathcal{K}} \rho_s m^I_{csk} r_{ck} \tag{12}$$

$$\text{subject to} \quad t_{cs} - \tilde{R}_s(0, a_c, n_c, p_c) - \sum_{k \in \mathcal{K}} m^R_{csk} r_{ck} = 0 \qquad c \in \mathcal{C}, s \in \mathcal{S}, \tag{13}$$

$$\sum_{c \in \mathcal{C}(g,h)} r_c - \sum_{c \in \mathcal{C}(g+1,h)} r_c \leq 0 \qquad \begin{aligned} & g \in \mathcal{G}, h \in \mathcal{H}, \\ & g < |\mathcal{G}|, \end{aligned} \tag{14}$$

$$\sum_{c \in \mathcal{C}(g)} v_c t_{cs} - L_g \sum_{c \in \mathcal{C}} v_c t_{cs} \geq 0 \qquad g \in \mathcal{G}, s \in \mathcal{S}, \tag{15}$$

$$\sum_{c \in \mathcal{C}(g)} v_c t_{cs} - U_g \sum_{c \in \mathcal{C}} v_c t_{cs} \leq 0 \qquad g \in \mathcal{G}, s \in \mathcal{S}. \tag{16}$$

The objective function (12) maximises the expected net interest income over all scenarios $s \in \mathcal{S}$ and customer segments $c \in \mathcal{C}$ by adjusting the income function $\tilde{I}(r_c, a_c, n_c, p_c)$ with the future response scenario $\tilde{R}_s(r_c, a_c, n_c, p_c)$ by means of a linear approximation. The values of the take-up variables $t_{cs}$ are approximated through the constraints (13). The constraint set (14) ensures monotonicity among the optimal rates for different risk categories. For example, if two segments $c_1 \in \mathcal{C}$ and $c_2 \in \mathcal{C}$ are similar with respect to their loan amounts and terms, *i.e.* $\mathcal{H}(c_1) = \mathcal{H}(c_2)$, but $c_1$ has a lower risk classification compared to $c_2$, *i.e.* $\mathcal{G}(c_1) \leq \mathcal{G}(c_2)$, then the inequality $r_{c_1} \leq r_{c_2}$ should hold. The constraint sets (15) and (16) imposes a lower and upper bound on the total volume for each risk grading $g \in \mathcal{G}$, over all the scenarios $s \in \mathcal{S}$.

## 5   Optimisation results

In order to illustrate the benefit of casting the retail credit price optimisation problem into a two-stage stochastic programming framework, empirical tests were performed using the data already introduced in § 2.

The response function $\tilde{R}_s(r, a, n, p)$ implies a set of betas being dependent on a scenario $s \in \mathcal{S}$. An approach to calculate a set of betas for each of the scenarios is by partitioning the

data set into subsets and estimating for each subset a response function. For the empirical results that will follow two data sets were created. For the first data set, referred to as the *expected scenario*, the data was not partitioned and a single response function was fitted to the entire segmented data set, *i.e.* $\mathcal{S} = \{1\}$ and the segment averages were used as input such that the number of cases were $|\mathcal{C}|$. For the second data set, referred to as the *repurchase rate scenarios*, three scenarios were created based on the partitioning done as illustrated in Figure 3. Note, however, that there is not an exact mapping between the three scenarios and the three time periods depicted in Figure 3. The first scenario corresponds to the sensitivity towards price due to a rapid decrease in repurchase rate, *i.e.* the first time period. The second scenario corresponds to the *third* time period and captures the sensitivity towards price due to a constant repurchase rate. Since our data set only includes historic periods where either a decrease in repurchase rate or a constant rate was observed, the third scenario was generated artificially to represent a potential future scenario during which increase in repurchase rate is expected that could result in potential customers being more price sensitive. This was achieved by modifying the regression coefficients of the second scenario's response function manually in order to make the take-up percentage much more sensitive towards price. The response graphs for the three repurchase rate scenarios that will be used as input to the SLPOP are shown in Figure 6.
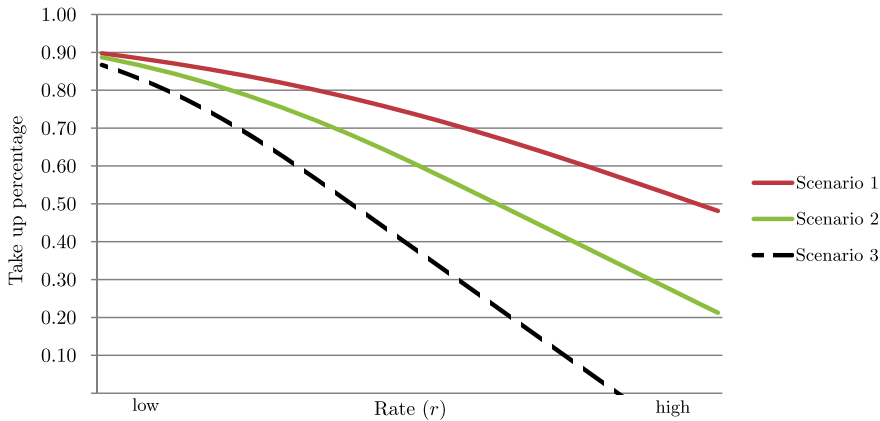


**Figure 6:**  *Response scenarios based on repurchase rate realisations.*

The SLPOP is formulated to cater for different customer segments in order to align with a customer centric pricing approach. Customer segments were created from the same historic data used for the estimation of the response functions and it was done in a fashion that will simplify practical pricing implementation. That is, pricing tables in practice entails having predefined prices for equally spaced intervals of the required loan amount, term and probability of default to simplify price lookup for a prospective client. With equally spaced intervals for each of the variables in the data set, each segment was then selected such that a unique combination of intervals is obtained for each of the variables. For example, let $\mathcal{B} = \{1, 2, \ldots, |\mathcal{B}|\}$ be the indices of the equally spaced intervals defined for the variables $a$, $n$ and $p$. The randomly selected sequence $\{b_a, b_n, b_p\}$ with $b_a, b_n, b_p \in \mathcal{B}$ forms one combination of interval indices which could be used to denote a segment. For

| Penalty | Average optimal prices | | | Take-up volumes | | |
|---|---|---|---|---|---|---|
| level ($\alpha$) | LR | MR | HR | LR | MR | HR |
| 0.00 | 59.7% | 64.4% | 67.5% | 30.7% | 27.1% | 42.2% |
| 0.50 | 53.1% | 57.0% | 60.5% | 30.7% | 26.9% | 42.4% |
| 1.00 | 46.3% | 50.0% | 53.5% | 31.0% | 26.9% | 42.0% |
| 1.50 | 39.4% | 42.8% | 46.5% | 30.7% | 26.8% | 42.4% |
| 2.00 | 33.0% | 35.8% | 39.3% | 31.0% | 26.4% | 42.6% |
| 2.50 | 25.7% | 28.6% | 32.3% | 31.0% | 27.0% | 42.0% |
| 2.75 | 22.7% | 25.0% | 28.8% | 31.0% | 26.5% | 42.6% |

**Table 1:** *Optimisation results showing the effect of the penalty level $\alpha$ on the SLPOP for the expected response scenario.*

testing purposes a total of 1016 segments were used in this study.

The numerical work in this study was performed by means of the SAS® software system [17]. The estimation of the response functions (6) and (11) was estimated with maximum likelihood estimation using the SAS® procedure *proc nlp*. An effective way to speed up the maximum likelihood estimation procedure is to make use of the beta solutions obtained from fitting the original response function as defined in (1) as starting solutions. This can be done effectively using the SAS® procedure *proc logistic*. The optimisation of the SLPOP was achieved through the use of the SAS® mathematical programming environment *proc optmodel*.

The subsections below address different aspects of solving the SLPOP. In the first subsection, the behaviour of the SLPOP as a result of using the expected response scenario is discussed. In the following subsection, the basic mechanism of the two stage stochastic programming framework is illustrated by solving the SLPOP for the three repurchase rate scenarios.

## 5.1 Solving SLPOP for the expected response scenario

The SLPOP formulation reduces to a deterministic optimisation problem when using the expected response scenario obtained from fitting the response model (11) to the entire data set. Initial results, however, showed that unrealistically high prices are generated as optimal solutions. This suggests that the fitted function $\tilde{R}_s(r, a, n, p)$ does not produce enough of a penalty effect on the net interest income. This could be attributed to the fact that the model in (11) does not take competitor information into account in representing price elasticity or that important variables may have been omitted. In an attempt to address this shortcoming the shifting parameter $\alpha$ in (11) is used to control the severity with which the response function $\tilde{R}_s(r, a, n, p)$ penalise the net interest income. The approach followed was to approximate the net interest income function (12) and the response function (13) of the SLPOP for different levels of $\alpha$ after fitting the model (11) with $\alpha = 0$.

Table 1 shows the results obtained by solving SLPOP for the expected scenario with different penalty levels of $\alpha$. The column "Average optimal prices" in Table 1 gives the optimal prices obtained for the risk gradings "LR", "MR" and "HR", which are risk categories defined for low risk, medium risk and high risk customers, respectively. Let $r_c^* = \sum_{k \in \mathcal{K}} r_{ck}^*$ denote the optimal price obtained for customer segment $c \in \mathcal{C}$ by solving

the SLPOP to optimality, then the average optimal price per risk grading $g \in \mathcal{G}$ is given by $(1/|\mathcal{C}(g)|) \sum_{c \in \mathcal{C}(g)} r_c^*$. The entries in the column "Take-up volumes" are calculated from the solutions to the take-up variables $t_{cs}^*$. That is, the percentage of take-up volume for risk grade $g \in \mathcal{G}$ relative to the total volume is given by $\sum_{c \in \mathcal{C}(g)} v_c t_{cs}^* / \sum_{c \in \mathcal{C}} v_c t_{cs}^*$, with $v_c$ the expected number of loan applications for customer segment $c \in \mathcal{C}$.

The adjustment effect of the response function due to adverse selection is clearly noticeable by observing that lower optimal prices are obtained for the low risk customers compared to the high risk customers for all levels of $\alpha$. Furthermore, the penalty effect on the net interest income can clearly be observed since optimal prices across all three risk categories are reduced with an increase in $\alpha$. The choice of an optimal $\alpha$ to use will be determined by how realistic the proposed optimal prices are. One way of doing this is to consider the legal limit imposed on pricing retail loans. Within the South African context the legal limit is given by the formula $2.2r_0 + 0.2$ with $r_0$ the repurchase rate. Considering that at the time of this study the repurchase rate was at 5%, the maximum allowable price legally is 31%. Therefore, the optimal $\alpha$ that will be used in the remainder of this paper will be $\alpha = 2.75$ since the optimal prices obtained are below the legal limit over all three risk categories. Limiting the prices could also be achieved by imposing an upper bound on the price variables $r_c$. However, the suggested penalty approach aims at controlling the severity of the response function while maintaining its functional form across different risk categories and different customer segments. This will not be achieved by only imposing an upper bound on the pricing variables. From the results in Table 1 it is clear that this objective was achieved since the volume distribution over the three risk gradings, given by the take-up volumes, maintained the same proportions for an increase in $\alpha$.

## 5.2 Solving SLPOP for the repurchase rate scenarios

In order to show the benefit of using a stochastic programming framework evidence must be provided that show an improvement in net present income by solving the SLPOP with different scenarios compared to solving the SLPOP using only the expected scenario (which reduces the SLPOP to a simple deterministic optimisation problem). To achieve this a test was performed by which optimal prices $r_c^{ES}$, for each customer segment $c \in \mathcal{C}$, are obtained by solving the SLPOP with the expected scenario (ES) and by re-calculating the objective function

$$\sum_{c \in \mathcal{C}} \sum_{s \in \mathcal{S}} \tilde{I}(r_c^{ES}, a_c, n_c, v_c, p_c, \delta) \tilde{R}_s(r_c^{ES}, a_c, n_c, p_c) \tag{17}$$

using the three repurchase rate scenarios. This was done to see if it would be reasonable to use the "expected" optimal prices in calculating the net present interest income for different realisations of future response scenarios. A desirable outcome, therefore, would be if the objective function value

$$\sum_{c \in \mathcal{C}} \sum_{s \in \mathcal{S}} \tilde{I}(r_c^{RSS}, a_c, n_c, v_c, p_c, \delta) \tilde{R}_s(r_c^{RSS}, a_c, n_c, p_c), \tag{18}$$

obtained by solving SLPOP over all three repurchase rate scenarios simultaneously is greater than the objective function value provided by objective function (17). In objective

| Future repurchase | Scenario probabilities | | |
| --- | --- | --- | --- |
| rate scenarios | (1) Decrease | (2) Constant | (3) Increase |
| Decreasing | 60.0% | 20.0% | 20.0% |
| Constant | 20.0% | 60.0% | 20.0% |
| Increasing | 20.0% | 20.0% | 60.0% |

**Table 2:** *Input cases based on different repurchase rate levels and future scenarios.*

function (18) $r_c^{RRS}$ are the optimal prices obtained by solving the SLPOP over all three repurchase rate scenarios (RRS) simultaneously. For subsequent empirical results the use of SLPOP requires the scenario probabilities $\rho_s$, for each scenario $s \in \mathcal{S}$.

Table 2 provides three cases that relate to different repurchase rate scenarios and arbitrary selected scenario probabilities. For the first case we assume that the current repurchase rate is high and a meaningful future outcome is that the repurchase rate will most likely decrease. For this case the probabilities of 60%, 20% and 20% were assigned to a decreasing, a constant and an increasing repurchase rate scenario respectively. For the case where the current repurchase rate is on an average historic level the probabilities of 20%, 60% and 20% were assigned to the three repurchase rate scenarios respectively, to indicate that we expect the repurchase rate to remain unchanged. For the final case a low current repurchase rate is assumed with the future expectation that it will increase, making it reasonable to assign the probabilities of 20%, 20% and 60% to a decreasing, a constant and an increasing repurchase rate scenario respectively.

Table 3 contains the optimisation results for the first case when considering a high current repurchase rate with a high probability that it may decrease in future. The first row provides the average optimal prices for the low risk (LR), medium risk (MR) and high risk (HR) categories respectively when solving SLPOP using all three repurchase rate scenarios simultaneously. For the same row the value in the column "Improvement" gives the percentage improvement in objective function value obtained in solving the SLPOP with all three repurchase rate scenarios, compared to solving the SLPOP using only the expected scenario. That is, the percentage improvement is obtained by calculating the relative improvement of (18) over (17). The improvement of 6.4% in net present interest income shows that the optimal prices obtained by solving the SLPOP over the three repurchase rate scenarios simultaneously are more robust compared to the expected optimal prices. More information is taken into account and optimal prices are better balanced against the effect of scenario dependent take-up volumes that influence the objective function value. In the first row of Table 3 take-up volumes for the HR category for the first scenario represents about 50% of the total volume. For the third scenario, however, the proportion of high risk volume is almost 80% making it a very risky portfolio. It could be argued that since there is only a 20% probability of the third scenario realising in future that it may be acceptable.

The SLPOP, however, does provide the capability to manage the portfolio risk by using the volume constraint sets (15) and (16). The second row of Table 3 gives the results for solving SLPOP using all three repurchase rate scenarios simultaneously, but with constraint set (16) imposing an upper bound of 50% on the high risk category volume. From the results it is clear that a price is paid in terms of the net interest income with the

| Data set | Average optimal prices | | | Improvement | Repurchase rate scenarios | Scenario probability | Realised take-up volumes | | |
|---|---|---|---|---|---|---|---|---|---|
| | LR | MR | HR | | | | LR | MR | HR |
| Repurchase rate scenarios | 17.9% | 21.1% | 25.9% | 6.4% | 1 (decrease) | 60% | 24.6% | 24.1% | 51.3% |
| | | | | | 2 (constant) | 20% | 40.0% | 31.1% | 28.9% |
| | | | | | 3 (increase) | 20% | 9.0% | 11.9% | 79.1% |
| Repurchase rate scenarios vol(HR) ≤ 50 | 16.8% | 19.5% | 26.4% | 5.3% | 1 (decrease) | 60% | 26.3% | 27.8% | 45.9% |
| | | | | | 2 (constant) | 20% | 41.3% | 37.1% | 21.6% |
| | | | | | 3 (increase) | 20% | 20.2% | 29.8% | 50.0% |

**Table 3:**  *Optimisation results for the case with an expected decrease in repurchase rate.*

objective improvement now only at 5.3%. However, the benefit achieved is an improved
risk profile with the high risk category volume limited to 50% for all three potential future
scenarios.

The optimisation results for the second case is shown in Table 4, where an average repur-
chase rate (compared historically) is considered with a high probability that it will remain
unchanged in future. Once again an improvement in objective function value is observed
when solving the SLPOP over the three repurchase rate scenarios simultaneously. The
first row shows an improvement of 10.6% but again with a risky portfolio if scenario three
happen to realise in future. The second row of Table 4 shows the results for resolving
SLPOP with the 50% limit on high risk volume with an improvement of 8.9% in net in-
terest income obtained over solving the SLPOP with the expected scenario. It should be
noted that, not only is an improvement achieved in net interest income by solving the
SLPOP over all three repurchase rate scenarios simultaneously, but it is also guaranteed
that the scenario dependent take-up volumes adheres to the volume constraint, which is
not necessarily the case for solving the SLPOP using the expected scenario.

| Data set | Average optimal prices | | | Improvement | Repurchase rate scenarios | Scenario probability | Realised take-up volumes | | |
|---|---|---|---|---|---|---|---|---|---|
| | LR | MR | HR | | | | LR | MR | HR |
| Repurchase rate scenarios | 17.3% | 20.0% | 24.0% | 10.6% | 1 (decrease) | 20% | 24.2% | 23.6% | 52.1% |
| | | | | | 2 (constant) | 60% | 38.2% | 29.6% | 32.2% |
| | | | | | 3 (increase) | 20% | 9.9% | 11.5% | 78.6% |
| Repurchase rate scenarios vol(HR) ≤ 50 | 15.8% | 17.7% | 25.1% | 8.9% | 1 (decrease) | 20% | 26.2% | 27.8% | 46.0% |
| | | | | | 2 (constant) | 60% | 40.4% | 36.6% | 23.0% |
| | | | | | 3 (increase) | 20% | 20.4% | 29.6% | 50.0% |

**Table 4:**  *Optimisation results for the case with the expectation that repurchase rate will remain
unchanged.*

The optimisation results for the third case is shown in Table 5. This scenario consists
of a low current repurchase rate with a high probability that there will be an increase in
repurchase rate in the future. The improvements obtained in solving the SLPOP over the
three repurchase rate scenarios simultaneously without and with the volume constraint are
20.9% and 20.2% respectively. These results show that meaningful results can be obtained
for the credit price optimisation problem by considering hypothetical scenarios like we did
with the inclusion of scenario three that represents a potential increase in repurchase rate,
a situation that is not captured in the historical data available.

| Data set | Average optimal prices | | | Improvement | Repurchase rate scenarios | Scenario probability | Realised take-up volumes | | |
|---|---|---|---|---|---|---|---|---|---|
| | LR | MR | HR | | | | LR | MR | HR |
| Repurchase rate scenarios | 16.1% | 19.6% | 24.9% | 20.9% | 1 (decrease) | 20% | 25.4% | 24.2% | 50.5% |
| | | | | | 2 (constant) | 20% | 39.6% | 30.4% | 30.0% |
| | | | | | 3 (increase) | 60% | 16.1% | 15.3% | 68.7% |
| Repurchase rate scenarios vol(HR) ≤ 50 | 14.6% | 18.2% | 25.6% | 20.2% | 1 (decrease) | 20% | 26.7% | 27.3% | 46.0% |
| | | | | | 2 (constant) | 20% | 40.1% | 35.6% | 23.5% |
| | | | | | 3 (increase) | 60% | 22.3% | 27.7% | 50.0% |

**Table 5:** *Optimisation results for the case with an expected increase in repurchase rate.*

# 6 Summary and conclusion

The price optimisation problem addressed in this paper deals with determining the optimal prices to quote prospective customers while considering uncertainty in future price sensitivity. That is, the take-up rates of future loans may deviate from current levels necessitating the use of a stochastic programming approach. This study is, to the best of our knowledge, a first attempt to incorporate uncertainty in price sensitivity as part of an explicitly formulated mathematical programming problem.

A concave response function is suggested that allows for the formulation of a linearised price optimisation problem that can be solved to optimality using standard linear programming technology. With a linear representation more complex formulations of the price optimisation problem can be handled such as volume constraints expressed in terms of the response function. The suggested response model also ensures a finite support for the pricing decision variable which makes it much more realistic compared to using a logit based response function with asymptotic properties.

The benefit of employing a stochastic programming approach was illustrated by means of empirical tests based on real data. Although some theoretical contributions have been made recently towards developing price optimisation models for retail credit, little evidence exist in literature of empirical work supporting the benefits of employing price optimisation technology. The results in this study showed that by only using expected optimal prices either a loss in revenue can occur due to lost opportunity, or a violation of strategic constraints may be expected when certain take-up scenarios realise in future.

# References

[1] AGRAWAL V & FERGUSON M, 2007, *Bid-response models for customised pricing*, Journal of Revenue and Pricing Management, **6(3)**, pp. 212–228.

[2] AUSUBEL L, 1999, *Adverse selection in the credit card market*, [Online], [Cited October 28th, 2013], Available from http://www.ausubel.com/creditcard-papers/adverse.pdf

[3] BOYD S & VANDENBERGHE L, 2004, *Convex Optimization*. Cambridge University Press, Cambridge.

[4] CAUFIELD S, 2012, *Consumer Credit Pricing*, pp. 138–152 in ÖZER Ö & PHILLIPS R (EDS), *The Oxford Handbook of Pricing Management*, Oxford University Press, Oxford.

[5] CROSS RG & DIXIT A, 2005, *Customer-centric pricing: The surprising secret for profitability*, Business Horizons, **48(6)**, pp. 483–491.

[6] EINAV L, JENKINS M & LEVIN J, 2012, *Contract pricing in consumer credit markets*, Journal of the Econometric Society, **80(4)**, pp. 1387–1432.

[7] EDELBERG W, 2006, *Risk-based pricing of interest rates for consumer loans*, Journal of Monetary Economics, **53(8)**, pp. 2283–2298.

[8] ENGELMAN B, 2011, *Measures of a rating's discriminative power: Applications and limitations*, pp. 269–291 in ENGELMAN B & RAUMEIER R (EDS), *The Basel II risk parameters*, Springer-Verlag, Berlin.

[9] HIGLE JL, 2005, *Stochastic programming: Optimization when uncertainty matters*, pp. 30–53 in GREENBERG HJ & SMITH JC (EDS), *Tutorials in Operations Research: Emerging theory, methods, and applications*, INFORMS, Hanover.

[10] KARLAN DS & ZINMAN J, 2008, *Credit elasticities in less-developed economies: Implications for microfinance*, American Economic Review, **98(3)**, pp. 1040–1068.

[11] OLIVER BV & OLIVER RM, 2014, *Optimal ROE loan pricing with or without adverse selection*, Journal of the Operational Research Society, **65(3)**, pp. 435–442.

[12] PARK S, 1997, *Effects of price competition in the credit card industry*, Economics Letters, **57(1)**, pp. 79–85.

[13] PHILLIPS RL, 2005, *Pricing and revenue optimization*. Stanford University Press, Stanford (CA).

[14] PHILLIPS RL & RAFFARD R, 2009, *Theory and empirical evidence for price-driven adverse selection in consumer lending*, In Proceedings of the 4th Credit Scoring Conference, Edinburgh, pp. 1–9.

[15] PHILLIPS RL, 2012, *Customized pricing*, pp. 465–490 in ÖZER Ö & PHILLIPS R (EDS), *The Oxford Handbook of Pricing Management*, Oxford University Press, Oxford.

[16] PHILLIPS RL, 2013, *Optimizing prices for consumer credit*, Journal of Revenue and Pricing Management, **12(4)**, pp. 360–377.

[17] SAS 9.1.3, 2011, *Help and documentation*, SAS Institute, Cary (NC).

[18] SKUGGE G, 2011, *The future of pricing: Outside-in*, Journal of Revenue and Pricing Management, **10(4)**, pp. 392–395.

[19] STIGLITZ J & WEISS A, 1981, *Credit rationing in markets with imperfect information*, American Economic Review, **71(3)**, pp. 393–410.

[20] THOMAS LC, 2009, *Consumer credit models: Pricing, profit and portfolios*. Oxford University Press, Oxford.