Title
Optimal ROE loan pricing with or without adverse selection

Journal
Journal of the Operational Research Society, 65(3)

ISSN
0160-5682

Authors
Oliver, BV
Oliver, RM

Publication Date
2014-03-01

DOI
10.1057/jors.2012.87

Peer reviewed

Journal of the Operational Research Society (2012), 1–8

www.palgrave-journals.com/jors/

Optimal ROE loan pricing with or without adverse selection

BV Oliver[1] and RM Oliver[2*]

[1]*Sandia Corporation, Albuquerque, USA; and* [2]*University of California, Berkeley, USA*

The authors describe the structural solution of the loan rate as a function of default and response risk that maximizes expected return on equity for a lender's portfolio of risky loans. Under the assumptions of our model, the non-linear differential equation for the optimizing price is found to be separable in transformed financial, response and risk variables. With an end-point condition where default-free borrowers are willing to borrow at loan rates higher than the lender's cost of funds, general solutions are obtained for cases where default probabilities may depend explicitly on the offered loan rate and where adverse selection may or may not be present. For the general solution, we suggest a numerical algorithm that involves the sequential solutions of two separate transcendental equations each one of which depends on parameters of the risk and response scores. For the special case where the borrower's default probability is conditionally independent of loan rate, it is shown that the optimal solution is independent of Basel regulations on equity capital.

Journal of the Operational Research Society advance online publication, 11 July 2012
doi:10.1057/jors.2012.87

Keywords: risk-based pricing; loan pricing; risk scores; response scores; banking; finance

Risk-based pricing

Risk-based pricing has developed considerable interest in recent years as a way of offering lenders a rationale for pricing loans to borrowers having different levels of default and response risk—see, for example, Oliver (2001), Phillips and Raffard (2009) and Thomas (2009). The simplest characterization of the problem suggests that the loan rate offered by a lender to a risky borrower should increase as the risk of default increases in order to cover the expected increase in default losses. At the same time, borrower preferences for lower prices must be recognized in developing a theory for the determination of 'optimal' prices for loan offers. One should have a reasonably clear idea of the objectives of the lender that may include volume of business and revenues, as well as expected return on equity (ROE). In this paper, we are only concerned with maximizing expected ROE.

When there is little or no competition in the marketplace, a lender or bank can set lending rates and terms of an offer in isolation of other lenders. Within the credit marketplace, lenders must recognize that borrowers are influenced by the presence and attractiveness of competitive offers. A full analysis of the interplay between many lenders and many borrowers would take us into multi-person and multi-objective game formulations where partial information between lenders and borrowers may be shared, much information is held privately and there is the opportunity for substantial lender–borrower negotiations before a loan is actually booked. Borrowers and lenders have asymmetric information and lenders may only be able to collect and process a limited amount of relevant risk data for individual borrowers. Although one can describe models where M borrowers continuously negotiate with N lenders to uncover optimal strategies or market equilibria, it is difficult to obtain insights for their solutions. Not only does each lender have different information about the risk and likely response of each borrower, but he or she must also assess the effect of cooperative or competitive policies that might be used. As offer terms and risk assessments change over time, the strategies of competing lenders may also change so that a model-builder soon becomes involved in dynamic formulations of substantial complexity that are unwieldy and yield little insight for the pricing process.

In this paper, we consider a single lender and the uncertain default risks and responses of multiple borrowers; the lender is a decision maker who structures and chooses an appropriate offer for each prospective borrower. Each borrower may accept the offer, go elsewhere or forego the loan. The lender's decision is a deterministic action taken against individual borrower risk and response uncertainty, whereas the decision of each borrower is viewed as the random outcome of an uncertain response to each lender offer. We assume that the lender has or can

*Correspondence: *RM Oliver, University of California, 260 Southampton Ave, Berkeley, CA 94707, USA.*
E-mails: oliver@ieor.Berkeley.edu; bvolive@sandia.gov

obtain as many well-calibrated risk and response scores about each borrower as are needed and that the objective is to maximize expected ROE. Although not explicitly discussed in this paper, mixed volume-ROE objectives are reasonably easy to model and solve.

Pricing for optimal ROE without default risk

Before we explicitly incorporate default risk, let us review the case where the lender attempts to find the best loan price offer, r, assuming the borrower Takes the offer with probability $q(r)$. In other words, the only uncertainty is the response to the offer, not the repayment of the loan. To make a sensible offer, we predict the influence of a loan rate decision on the borrower's preferences and how the probability of response influences the lender's expected ROE. This pricing problem is familiar to experts in marketing and yield management, and focuses on price as a decision variable to control profit rather than a way to control risk. We now know that price simultaneously influences both risk and profit.

We assume that the cost of borrowed funds to source the loan is c_D; if an offer is not Taken by the prospective borrower, the available but 'unused' equity capital, E, is invested in risk-free securities at the risk-free rate r_F. We also assume that the commercial borrowing rate for the lender is greater than the risk-free rate but less than the loan rate for the retail borrower, as it would not be profitable for the lender to borrow funds at one rate and then loan it to an individual borrower at a lower rate. As we are silent about the possibility of default, the expectation of ROE is

$$\mathbb{E}[r_E(r)] = r_F + \frac{1}{E}(r - c_D)q(r) \qquad r > c_D > r_F \tag{1}$$

Although the risk-free rate enters the expression for expected ROE it does not affect the loan rate decision. It is easy to show that the optimal lending rate is independent of r_F and must satisfy the first-order differential equation:

$$q^* + (r - c_D)\left.\frac{dq}{dr}\right|_{r=r^*} = 0 \qquad q^* = q(r^*) \tag{2}$$

With small changes in the lender's offer rate, the first term on the left-hand side represents the additional expected volume that would be obtained at the optimal booking rate; because the derivative of the Take rate is negative, the second term represents the net profit that results from the decrease in volume caused by small changes in loan rate. At the optimum these terms offset. The optimal solution for r can also be expressed in terms of a price-response elasticity, ε. At the optimum the profit from existing borrowers lost by decreasing the price is compensated by the increased demand from new borrowers so that the elasticity at the optimal price is

$$\begin{aligned} \varepsilon^* \triangleq \left.\frac{dq}{q}\frac{r}{dr}\right|_{r=r^*} = \frac{r^*}{c_D - r^*} < -1 \\ r^* > c_D \Leftrightarrow -\infty < \varepsilon^* = \varepsilon(r^*) < -1 \end{aligned} \tag{3}$$

Our definition of response elasticity yields negative values because we want to emphasize that response or demand decreases with increases in loan rate or price—the customary definition in the economics literature is the negative of our elasticity. It is clear from the assumptions of the model that the optimal loan rate determined by this solution is not risk dependent and cannot, in its present form, recognize the effect of adverse selection. Rather it provides a single 'optimal' loan rate for certain repayment. The necessary condition for an optimal solution in (3) can be rewritten as:

$$\frac{r^* - c_D}{r^*} = -\frac{1}{\varepsilon^*}$$

The term on the left-hand side is called the *premium or contribution margin ratio*. It is the premium of return of the loan rate over the borrowing rate in units of loan rate and is often referred to as the inverse elasticity rule (Wilson, 1993; Phillips, 2005). Alternatively, the optimal loan rate can be expressed in terms of an elasticity factor times the borrowing rate:

$$r^* = \frac{\varepsilon(r^*)}{1 + \varepsilon(r^*)} c_D > c_D \qquad \varepsilon(r^*) < -1 \tag{4}$$

The mathematical structure of this formula tells us that the optimal lending rate is the cost of borrowed funds multiplied by a positive elasticity factor that depends on the attractiveness of marginal increases in loan rate to the retail borrower. The same result can be obtained when elasticities are defined in terms of the premium between lending and borrowing rates. In retail lending, the factor that includes the elasticity terms is >1, which means that the optimal retail loan rate is greater than the cost of commercial borrowed funds. For example, if the elasticity is -1.5, (4) tells us that $r^* = 3c_D$; if -2, $r^* = 2c_D$; if -3, $r^* = 1.5c_D$. Because both sides of the equation depend on the optimal offer rate, it is clear that we must solve an implicit equation in which the rate used to calculate the elasticity coincides with the optimal rate. This observation also suggests that experimental estimates of response elasticity should be made in the vicinity of the optimal loan offer. In general it is not possible to obtain analytical solutions. The family of exponential and linear response rates is an exception but, in our experience, these solutions are hardly ever useful in business situations. As we find in later sections, the elegant simplicity and structure of (4) does not appear to carry over to those cases where both price-induced response and price-induced risk elasticities appear side by side.

The optimal premium due to leveraged investment of the equity E is therefore

$$\mathbb{E}[r_E(r^*)] - r_F = \frac{1}{E}(r^* - c_D)q(r^*) = \frac{-c_D q(r^*)}{E(1+\varepsilon^*)}$$
$$\varepsilon^* = \varepsilon(r^*) < -1, \qquad (5)$$

which depends on the magnitude of the leverage, the cost of funds borrowed by the lender, the Take rate and the elasticity at the optimal price. If the elasticity factor in (5) is slightly smaller than -1 and the equity E (capital reserve) required to borrow funds for a unit loan is a fraction, say 10%, the denominator is small and the expected ROE on loans can become very large as it is inversely proportional to the equity level. Because the attractiveness of offers to individual borrowers affects the fraction that book, the expected return on the portfolio depends on response elasticity of borrowers and prices.

Although there has been no explicit inclusion of borrower default, one significant risk to the lender comes from the competitive marketplace and the unfulfilled promise of individual borrowers to repay their loans. In retail and commercial lending applications, default risk must be explicitly included in a pricing decision as there may be large losses from accounts that default. ROE should therefore include losses from uncertain defaults on assets, as well as the expense of debt on the liabilities of the balance sheet.

Optimal ROE pricing for a loan portfolio

We now consider the acquisition of many different borrowers, each with their own risk of default and response to an offer r; we again use the notation E for equity capital, r the lender's loan rate and revenue from the borrower of a unit loan that does not default, c_D the lender's cost of borrowed funds that source the loan and l_D for the lender's loss if the loan defaults, usually referred to as LGD (loss given default). In deriving a mathematical expression for the expectation of the random ROE, we note that the net return for a lender who borrows money to source the loan is $r-c_D$ for the Good (no default) borrower and $-(l_D+c_D)$ for a Bad (default) borrower. In the case of default, the lender must pay l_D+c_D because obligations for borrowed funds require full repayment with interest—that is, there is no relief to the debt of a commercial lender or bank because of the default of a retail borrower whose loan was sourced by the borrowed funds. We denote the conditional probability of non-default (Good, G) and default (Bad) for a booked borrower who has taken the lender's offer

$$p(G|T,\mathbf{x},r) = \Pr\{Good|Take,\mathbf{x},r\} \triangleq p(\mathbf{x},r)$$
$$p(B|T,\mathbf{x},r) = \Pr\{Bad|Take,\mathbf{x},r\} = 1 - p(\mathbf{x},r) \qquad (6)$$

and the conditional probability of taking the offer as

$$\Pr\{Take|\mathbf{x},r\} = p(T|\mathbf{x},r) = q(\mathbf{x},r) \qquad (7)$$

The unconditional expected return for the lender's portfolio (risky loans and risk-free assets) is the risk-free yield for equity plus a premium for the loans to the portfolio of risky borrowers who book at the risk-and response-dependent loan rate offered to them. The expected yield premium (above the risk-free rate r_F) for the portfolio of loans can be written as:

$$\begin{aligned}\mathbb{E}[r_E(r)] - r_F &= \frac{1}{E}\int_{\mathbf{x}\in X}\begin{pmatrix}(r-c_D)p(G|T,\mathbf{x},r)\\ -(l_D+c_D)p(B|T,\mathbf{x},r)\end{pmatrix}\\ &\quad\times p(T|\mathbf{x},r)dF(\mathbf{x}) \quad r\in R\\ &= \frac{1}{E}\int_{\mathbf{x}\in X}\begin{pmatrix}(r-c_D)p(\mathbf{x},r)\\ -(l_D+c_D)(1-p(\mathbf{x},r))\end{pmatrix}\\ &\quad\times q(\mathbf{x},r)dF(\mathbf{x}), \qquad (8)\end{aligned}$$

where $F(\mathbf{x})$ denotes the profile of borrowers acceptable for offers and hence for the booked portfolio. The interaction and tradeoffs between the Take probability and the loan rate are critical: if too large a rate is offered few borrowers will take the offer, whereas if the rate is too low the offer may appeal to many borrowers but may not be profitable for the lender. By lowering loan prices, market share increases but expected returns from booked accounts are not large enough to cover the borrowing costs and increased default losses. Because of the presence of two uncertainties (risk and response), attention shifts from finding risk cutoff policies that acquire customer portfolios with acceptable levels of risk to finding pricing policies that achieve expected profit, volume and risk objectives in an environment where a lender may be forced to compete with other lenders attracting the same risky borrowers. Surprisingly, some optimal prices also suggest cutoffs.

To reiterate the point made earlier a lender wants to find the loan rate, $r(p,q)$ that maximizes the expected premium provided by the risky investments over the risk-free ones. Without restrictions on regulatory capital or self-imposed volume requirements, the ROE objective is

$$\operatorname*{Max}_{r(p,q)\in R}(\mathbb{E}[r_E(r)] - r_F) \qquad (9)$$

The optimum is achieved by selecting an appropriate functional $r(p,q) = r(p(\mathbf{x},r), q(\mathbf{x},r))$ with the clear understanding that the risk and response probabilities are themselves functions of the loan rate and a vector $\mathbf{x}$ that characterizes the risk or preferences of an individual borrower. The subset of $\mathbf{x}$ that carries the relevant information for $p(\mathbf{x},r)$, the probability of non-default, is usually very different from the subset used in the prediction of response, $q(\mathbf{x}, r)$. For this reason one can seldom use the simpler relationship $q(p, r) = q(p(\mathbf{x}, r), r)$ to formulate (8).

The role of adverse selection should be clarified as it may or may not be present. In either case, the appropriate conditional risk probability is the posterior one associated with borrowers who Take the offer, and, because of private information available to themselves or the markets but not the lenders, may result in a different probability of default than would be expected from those who do not Take the offer, that is, the prior and posterior probabilities of Good/Bad for Takes and Non-Takes differ. Because the risk profile of booked loans is conditional on borrower preferences and the probability of a default is conditional on the loan being booked, the appropriate default probabilities are the ones that apply to the booked sub-population where the condition on a Take, T, is explicitly recognized. If there is no Adverse Selection then the posterior and prior probabilities of default are equal and independent of the act of booking. If there is adverse selection, the revised (posterior) probabilities for default given a Take should be used in which case their effects are naturally and correctly incorporated in the decision problem. For further details on Bad adverse selection and Good positive selection, see Oliver and Thaker (2012).

The formulation in (9) is a bit puzzling in that the integration (summation in discrete models) obviously removes the dependence on p, yet the maximization operator outside the integral asks us to find the best risk-dependent offer rate as a function of p and q. We emphasize that the decision is the risk—and response-dependent loan rate or price, whose mathematical structure we seek as a *function* of our assessments of these risks—the decision problem is no longer one of selecting an optimal operating point or cutoff even though a cutoff may be the natural consequence of some optimal pricing functions. Depending on the competitive marketplace, the offered loan rate directly affects borrower preferences, and acceptance by the borrower directly influences the composition of loan assets and debt obligations undertaken by the lender to source the risky loan portfolio.

Assume that $r(p, q)$ is continuous and has continuous first derivatives. The equation for the optimizing risk- and response-based price is obtained by setting the partial derivative of the expected ROE with respect to lending rate equal to 0. The profile for Good or Bad in the booked population is always non-negative. The necessary condition for a relative maximum is a first-order non-linear differential equation:

$$\frac{\partial}{\partial r} q(\mathbf{x}, r)((r - c_D)p(\mathbf{x}, r) - (l_D + c_D)(1 - p(\mathbf{x}, r)) = 0 \quad (10)$$

Typically, one provides end-point conditions when non-default of the borrower is certain but, in this problem, an alternative is to specify the desired value for expected unleveraged ROE premium. Notwithstanding the oft-used argument for perfect arbitrage in efficient markets, it should be noted that even when non-default is certain a lender may, nevertheless, get limited acceptance to an offer whose loan rate is higher than the risk-free rate. There is ample evidence that individual borrowers will pay high prices for convenience, a relationship with a lender or, more importantly, lack of information on the availability of better offers in the marketplace.

To simplify notation in (10), we substitute p for the conditional probability of a Good given a Take and q for the probability of a Take given an offer so that the differential equation in (10) can be rewritten as:

$$(r + l_D)\frac{\partial(pq)}{\partial r} - (l_D + c_D)\frac{\partial q}{\partial r} + pq = 0 \quad (11)$$

If the middle term (derivative of q) were not present the solution would be separable in two factors, one depending only on p, the other only on q. Fortunately, by using the simple translation

$$r' = r + l_D, \quad c'_D = c_D + l_D, \quad (12)$$

separability can be ensured and thus provide a considerable simplification in the solution. Using (12), (11) can be rewritten in terms of primed rates as:

$$\frac{\partial(pqr')}{\partial r'} = c'_D \frac{\partial q}{\partial r'} \quad (13)$$

The general solution is surprisingly simple:

$$r'p(\mathbf{x}, r')q(\mathbf{x}, r') = c'_D q(\mathbf{x}, r') + C \quad (14)$$

In this expression C is a constant of integration whose explicit value is independent of the three quantities r', $p(r')$ and $q(r')$; it can be determined from end-point conditions on a guaranteed no-default loan or, equivalently, a requirement on desired expected unleveraged ROE. Returning to the unprimed notation it is easy to show that the mathematical structure of the optimal loan rate is

$$\begin{aligned} r^*(p, q) &= r^*(p(\mathbf{x}, r^*), q(\mathbf{x}, r^*)) \\ &= \left(\frac{c_D}{p(\mathbf{x}, r^*)} + l_D \frac{1 - p(\mathbf{x}, r^*)}{p(\mathbf{x}, r^*)}\right) \\ &\quad + \left(\frac{C}{p(\mathbf{x}, r^*)q(\mathbf{x}, r^*)}\right) \qquad 0 < p, q \leqslant 1 \quad (15) \end{aligned}$$

We emphasize that conditional probabilities must be evaluated at the optimal r^*. Default and response risk probabilities in the second term are additive to the lender's risk-neutral price in the first term on the right-hand side, which means that the lender's risk-neutral solution provides a lower bound for the optimal price as $p(\mathbf{x}, r^*)q(\mathbf{x}, r^*)$ product is positive and <1. This solution yields the anticipated result that there is equality of optimal expected

ROE in all p, q, r combinations and that the constant of integration, C, is the unleveraged optimal ROE premium for the entire portfolio. The latter can be shown by substituting (15) into (8) and completing the integration. Although it may be possible to express the general solution in terms of price-risk and price-response elasticities, it is not clear to us how one can duplicate the elegant structure in (4).

The structure of the optimal loan rate in (15) is illustrated in the three-dimensional surface of Figure 1, where p, q and r axes are identified by the labels. All points on this surface yield the same expected ROE premium but there is no guarantee that an arbitrarily chosen point on the surface can satisfy a required theoretical or experimental relationship between p, q and r. Once we separately specify the mathematical structure of $p(\mathbf{x}, r)$ and $q(\mathbf{x}, r)$ and compute their inverses we can illustrate how default and response surfaces intersect the p, q, r surface in Figure 1 and provide a sub-set of solutions.

Ability to pay: default risk conditionally independent of loan price

When the ability of the borrower to repay is not in question we can remove the explicit dependence of the probability of default on loan rate in (13) by setting the derivative of p with respect to r equal to 0. Thus, the simpler differential equation for the first-order condition is now

$$\frac{\partial(r'q)}{\partial r'} = \frac{c'_D}{p}\frac{\partial q}{\partial r'} \tag{16}$$

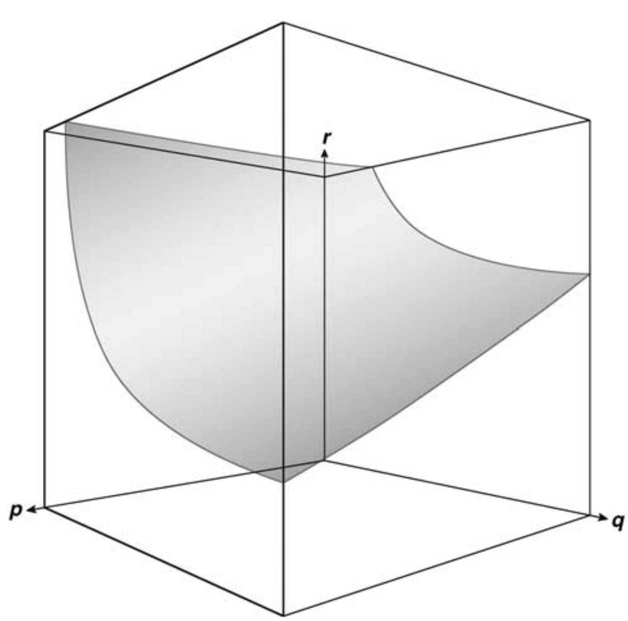

Figure 1 Constant optimal ROE surface as a function of (p, q, r).

In the integration of (16) we treat p as a constant that, nevertheless, can vary over the range of risk levels acceptable to the lender. Even when conditional independence is not assumed in the more general case of (11), a borrower's probability of default is weakly affected by the loan rate for large p (eg prime paper); one can also show mathematically that terms proportional to $\partial p/\partial r$ are small for all r close to c_D. For each fixed p the general loan pricing solution is now

$$r' = \frac{c'_D}{p} + \frac{C}{pq} \tag{17}$$

As before, the solution requires a meaningful end-point condition that fixes the loan rate when the probability of default is 0 and repayment is certain, that is $p=1$, but where attractiveness of the offer to the borrower may still be uncertain, that is $0<q<1$. Thus, in this case, the correction term to the familiar risk-neutral price of the lender includes only the marketing effect of a risk-based price for each risky borrower:

$$r^*(q(r^*)|p) = \left(\frac{c_D}{p} + l_D\frac{1-p}{p}\right) + \left(\frac{C}{pq(\mathbf{x}, r^*)}\right) \tag{18}$$

Compare this result with (15) where default depends explicitly on the offer rate. What is interesting about (18) is that Basel or other capital reserve requirements that depend on the borrower's probability of default do not affect the optimal solutions, even when equity capital E in (8) is replaced inside the integral by a risk-dependent $E(p)$. Because p is fixed and independent of r, equity E is also independent of r.

Consider the special case of optimal pricing for expected return on assets (ROA) in the absence of affordability issues. The lender does not borrow ($c_D=0$) and excludes the leveraging effect of equity by setting $E=1$. If we consider the watershed case where the risk-neutral lender wants to achieve an expected ROA equal to the risk-free rate, we set $C=r_F$ so that the optimal price in (18) becomes

$$r^*(q(\mathbf{x}, r^*)|p) = \frac{r_F}{pq(\mathbf{x}, r^*)} + l_D\frac{1-p}{p} \tag{19}$$

Simply stated: multiply the LGD by odds of a Bad to cover expected default losses and add to it the risk-free rate discounted by the product of the uncertainties. This result in (19) can be compared with Equation (3.4.8) in Thomas (2009). Multiplying (18) or (19) by the probability of Good/Bad outcomes for those who Take gives the desired ROA for each account, as well as the booked portfolio.

A referee has correctly pointed out that the necessary condition in (10) applies to each prospective individual borrower in a continuum of rates, preferences and default risks. What should the lender do in the case where prices are being offered in a small, finite number of segments or price/risk tiers and the alternatives represent choices in two or more discrete dimensions? If the overall objective is to maximize expected ROE, (15) tells us that without additional risk constraints the optimal expected ROE of all segments must be equal; thus, a discrete formulation should simultaneously balance tradeoffs of risk, borrower preferences and volume within each tier or segment! One can think of (15) as the defining price-risk-response relationship for the 'centre' of the tier segment and use it to design the boundaries rather than the other way round where tier boundaries are arbitrarily specified and an optimal pricing relationship is then sought. In other words, use the solutions of the continuous model to guide one's thoughts about discrete price, risk and response tiers. If, on the other hand, the tiers are predetermined and the operational situation requires a discrete formulation, then one should start at the beginning and optimize an objective (with appropriate monotinicity and end-point conditions) that represents the counterpart to (8) expressed in terms of discretized p_i, q_j, r_k and financial parameters.

Action-based response and risk scores

It is common practice to use risk and response scores to make predictions of default and whether a potential borrower may or may not Take a lender's offer. In most scoring applications the characteristics that play a part in the assessments are a mixture of discrete and continuous variables. In the procedures that follow we assume risk and response scores are linearly dependent on the continuous decision variable, r, whereas the behavioural and demographic characteristics in the $\mathbf{x}$ vector may be a mixture of continuous and discrete variables. In (20) and (21) below, $\beta_{(\bullet)}$ is a scalar slope representing the change in score per unit change in loan rate; a discrete version can be easily represented by a vector $\boldsymbol{\beta}_{(\bullet)}$ in the event one uses a finite segmentation of loan rates. The action-based response score for Takes and Non-Takes can be expressed in log odds form as:

$$s_q(\mathbf{x}, r) = \alpha_q(\mathbf{x}) - \beta_q r \qquad \beta_q > 0 \tag{20}$$

and the action-based default risk score has the general form:

$$s_p(\mathbf{x}, r) = \alpha_p(\mathbf{x}) - \beta_p r \quad \beta_p > 0. \tag{21}$$

We introduce a minus sign to emphasize the feature, observed experimentally, that both scores decrease with increasing r; this is not to say that the action-based scores in (20) and (21) are always less than baseline scores that would have been obtained in the absence of r. It should be emphasized that the intercept term $\alpha_{(\bullet)}(\mathbf{x})$, which is different for each individual borrower, is not itself a calibrated log odds score; the value provided by (21) may be larger or smaller than the baseline score. For further details see text following Equation (34) in Oliver and Thaker (2012). The probability of a Take is given by

$$q(\mathbf{x}|r) = \Pr\{Take|\mathbf{x}, r\} = \left(1 + e^{-s_q(\mathbf{x},r)}\right)^{-1} \tag{22}$$

and the probability of non-default is given by

$$p(\mathbf{x}|r) = \Pr\{Good|\mathbf{x}, r\} = \left(1 + e^{-s_p(\mathbf{x},r)}\right)^{-1} \tag{23}$$

An important reminder is that the forecast horizon for the predictions in (22) and (23) must be synchronized with the period used to describe the interest rates, economic returns and the ROE model in (8). We have assumed in our models that rates, returns and predictions of default are described in units of a common period, say a year. If the analysis uses probabilities based on a 2-year horizon, for example, then one must appropriately scale ROE rates and yields.

Numerical solutions in terms of scores

In solving the implicit equations in (15) or (18) there may be uninteresting mathematical solutions whose relevance to realistic lending rates has to be argued. For example, when we use log-odds risk and response score models there are occasions where one of two roots must be discarded on the grounds that the larger root represents a predatory loan rate or where the probability of response is so small that it leads to an inefficient solution with the same optimal ROE but lower probability of Take.

In the first model we studied in (1) there is no need for a risk score as there is no default risk. On substituting the Take probability in (22) into (4) the optimal loan rate is seen to be the solution of a transcendental equation in z

$$z = a + be^z, \tag{24}$$

where $z \triangleq \beta_q r, a = \beta_q(c_D + C), b = \beta_q Ce^{-\alpha_q}$. C has the interpretation of the additional interest rate (basis points) that a lender must add to his borrowing cost to fund a loan with a borrower who is certain to honour his loan contract and, at the same time, represents the expected unleveraged ROE premium over the risk-free rate. With a cost of funds equal to 3%, $E = 0.08$, $C = 0.025$, $\alpha_q = 3.5$, $\beta_q = 30$ (for the response score parameters) we find that the solution for the optimal loan rate is $r^* = 0.0595$, $q(r^*) = 0.822$ and $\mathbb{E}[r_E(r^*)] - r_F = 0.313$. Note that $E = 0.08$ (equity capital as a fraction of funds borrowed) corresponds to a leverage ratio of 12.5 which, with the optimal loan rate, yields an expected ROE about 10 times the 3% cost of funds.

A more important case occurs when the borrower has the ability to repay the loan obligations under any loan rate offer; the risk score and the probability of default are, a priori, dependent on behavioural and demographic data but independent of r, although, a posteriori to a borrower's response, the probability of default may implicitly depend on loan rate. Substituting the assumptions of (20) and (22) into (18) yields

$$r(p,q) = \left(\frac{c_D}{p} + l_D\frac{1-p}{p}\right) + \frac{C}{p}\left(1 + e^{-(\alpha_q(\mathbf{x})-\beta_q r)}\right) \quad (25)$$

The solution for optimal r with p fixed is again a transcendental equation of the form $z = a + bc^z$, where

$$z = \beta_q r, \quad a = \frac{\beta_q}{p}(C + c_D + l_D(1-p)), \quad b = \frac{C\beta_q}{p}e^{-\alpha_q} \quad (26)$$

As discussed earlier, p can assume all values representing acceptable default risks for the lender, for example all values above the lender's preliminary Accept/Reject cutoff. In (25) the probability of a Good, p, is conditionally independent of r so that an optimal solution, r^*, does not have a feedback effect that then alters the chosen p value. Finding a root is straightforward because only one transcendental equation is solved.

Figure 2 is a three-dimensional plot of the same ROE surface shown in Figure 1 but where two planes for two different fixed values of the probability of non-default (p-axis) and the inverse of a log odds score ribbon for the probability of a Take (q-axis) in (25) are superimposed on that surface. As in Figure 1, all (p, q, r) values increase in the direction of the arrows; the range of values in p are (0.85, 1), in q are $(0, 1)$ and in r are $(0, 0.25)$ with the origin having the value $(0.85, 0, 0)$ at the rear of the three-dimensional cube. The loan rate is the vertical axis and numerical values of (p, q, r) are identified on selected corners of the three-dimensional cube. The intersection of the constant ROE surface in Figure 1 with the $r-q$ ribbon (from (20) and (22)) results in a two-dimensional curve, denoted by ABC. This ABC curve is denoted by a thicker line extending behind the first semi-transparent $p = 0.97$ plane. This plane represents a default rate of 3% and has two intersections (roots) with the ABC curve. Both roots provide the same expected ROE but the larger root for r^* provides the less efficient solution with a smaller, almost negligible, volume and the smaller root corresponds to a higher q and larger expected volume. As p is reduced the two roots get closer and closer together on ABC until, finally, they coalesce into a single root that defines a point of tangency and 'risk cutoff' at B with the second, opaque $p = 0.94$ plane. Planes representing still smaller values of p, and higher default rates, have no intersection with ABC and hence no solution satisfying the required end-point condition at $p = 1$ or the C value desired for optimal expected ROE. In summary, the surface in Figure 1 representing (15) is independent of borrower preferences whereas the response ribbon provides a risk cutoff at point B that differs with each individual borrower. With optimal pricing the notion of a global risk cutoff for all borrowers may no longer be valid. When we examine values of $p = 0.99, 0.97, 0.94$ (default rates of 1, 3 and 6%, respectively) with response score parameters $\alpha_q = 3.5$, $\beta_q = 30$, we find that the loan rates maximizing expected ROE are $r^* = 0.066$, 0.081 and 0.119, respectively. These compare with the earlier example for the optimal loan rate of 5.95% when there is certainty in repayment.

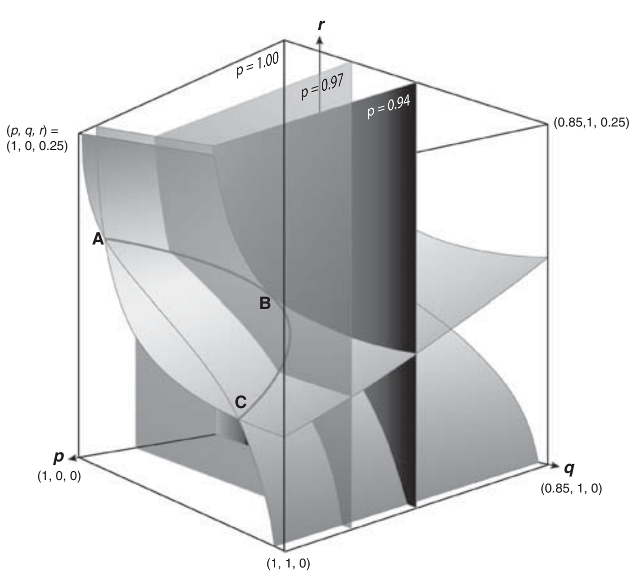

Figure 2 Optimal loan rate r with 3 and 6% default probabilities.

Numerical solutions are more difficult to visualize when both risk and response probabilities in (15) depend on the loan rate, r. Fortunately, one can state the general solution in terms of the parameters of the log odds scores in (20) and (21) and thus obtain a considerable simplification in calculating numerical values for r^*. This yields an optimal loan rate that is a solution of the triple exponential transcendental equation

$$f(r^*) = -r^* + C_0 + C_1 e^{\beta_p r^*} + C_2 e^{\beta_q r^*} + C_3 e^{(\beta_p+\beta_q)r^*} = 0, \quad (27)$$

where the constant of integration is C (without a subscript) and the four coefficients C_i, with subscripts $i = 0,1,2,3$, depend on C, the score intercept terms (α_p and α_q), the slopes (β_p and β_q) of the response and risk scores in (20) and (21), as well as the financial parameters for the commercial borrowing rate, c_D, and the LGD, l_D:

$$C_0 \triangleq C + c_D, \quad C_1 \triangleq (C + c_D + l_D)e^{-\alpha_p}, \quad C_2 \triangleq Ce^{-\alpha_q}, \quad C_3 \triangleq Ce^{-(\alpha_p+\alpha_q)} \quad (28)$$

Analytic closed-form solutions can be obtained for a few uninteresting special cases but not, as far as we know, for the score-based formulations that are used in the credit industry. Because we know that the p (risk) and q (response) functions can be specified in terms of their scores, optimal loan rate solution(s), when they exist, correspond to the intersection of three distinct surfaces. For each individual borrower specified by $(\mathbf{x}, r)$ there are two ribbons (surfaces) represented by the inverse functions $\bar{r} = p^{-1}(\mathbf{x}, \bar{p})$ and $\bar{r} = q^{-1}(\mathbf{x}, \bar{q})$ that may or may not intersect (15).

We have found that numerical solutions of (27) are considerably simplified by designing an algorithm that sequentially solves two distinct single-exponential transcendental equations in r. Solutions for the optimal loan rate are obtained by recursively solving two equations of the form of (24), each with its own set of parameters that correspond to the fixed p and fixed q cases described earlier. First, $p(r)$ is fixed and the coefficients of one transcendental equation are expressed in terms of finding a q-dependent root as shown in Figure 2. This new value of r yields a value of $q(r)$, which is then held fixed while a new p-dependent root for r is sought; the iterative process is repeated until there is convergence in the root. In our experience five decimal root accuracy (1 basis point in the interest rate) is obtained with a small number of iterations (<10) of each transcendental equation. As an illustrative example for the case where both risk and response surfaces are derived from log odds scores, we use data from previous examples except that log odds risk scores are substituted for fixed p values. With risk scores parameters given by $\alpha_p = 3.5$ and $\beta_p = 2$, we obtain $r^* = 0.085$, $p(\mathbf{x}, r^*) = 0.966$, $q(\mathbf{x}, r^*) = 0.722$ for loan rate, default and response probabilities.

It should be noted that the scorecard-based calculations require extensive computations, which, besides the preparation of the two scorecards and the posting of four additional column numbers opposite each record, require the solution of (27) and (28) for each borrower. What might appear to be a formidable and intractable set of calculations requiring extensive databases is well within the reach of existing high-speed computation platforms.

We appreciate the thoughtful critiques of both referees who encouraged us to improve our graphical presentations, the geometrical interpretation of solutions and insights on pricing effects on both ROE and lender preferences for more efficient choices.

Acknowledgements—This paper is dedicated to William R. Fair, one of the two founders of The Fair Isaac Corporation, a dear friend, and a person whose insights and encouragement ultimately led to this paper

References

Oliver RM (2001). Risk-based pricing. Paper presented at Edinburgh Credit Scoring and Credit Control Conference VII.

Oliver RM and Thaker A (2012). Adverse selection and non-take inference with coherent risk and response scoring. *Journal of the Operational Research Society*, advance online publication, 11 April, doi:10.1057/jors.2012.3.

Phillips R (2005). *Pricing and Revenue Optimization*. Stanford University Press: Palo Alto, CA.

Phillips R and Raffard R (2009). Theory and empirical evidence for price-driven adverse selection in consumer lending. Paper presented at Edinburgh Credit Scoring and Credit Control Conference XI, Revised as Columbia University Business School Working Paper in 20 July 2010.

Thomas LC (2009). *Consumer Credit Models*. Oxford University Press: New York.

Wilson RB (1993). *Non-Linear Pricing*. Oxford University Press: Oxford, UK.

Received January 2012;
accepted May 2012 after one revision