Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Banking & Finance 29 (2005) 1213–1236

Journal of BANKING & FINANCE

www.elsevier.com/locate/econbase

The relationship between default prediction and lending profits: Integrating ROC analysis and loan pricing ☆

Roger M. Stein

Research and Analytics, Moody's KMV, 99 Church Street, New York, NY 10007, USA

Received 20 August 2003; accepted 5 April 2004
Available online 3 September 2004

Abstract

In evaluating credit risk models, it is common to use metrics such as power curves and their associated statistics. However, power curves are not necessarily easily linked intuitively to common lending practices. Bankers often request a specific rule for defining a *cut-off* above which credit will be granted and below which it will be denied. In this paper we provide some quantitative insight into how such cut-offs can be developed. This framework accommodates real-world complications (e.g., "relationship" clients). We show that the simple cut-off approach can be extended to a more complete pricing approach that is more flexible and more profitable. We demonstrate that in general more powerful models are more profitable than weaker ones and we provide a simulation example. We also report results of another study that conservatively concludes a mid-sized bank might generate additional profits on the order of about $4.8 million per year after adopting a moderately more powerful model.

JEL classification: C52; C25; C14; C15; C44
Keywords: Default prediction; ROC analysis; Model validation; Rating systems; Economic value of information; Discrete choice models; Model power; Simulation; Bank lending

☆ An earlier draft of this paper appeared as a working paper with the title "Power, Profitability and Prices".

E-mail address: steinr@moodys.com

0378-4266/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbankfin.2004.04.008

1. Introduction

The use of credit scoring models by financial institutions has increased dramatically. With this increase has come a need among users to understand the economic value of the models and to use this information to integrate them into traditional lending practices in a profitable manner. For example, many institutions find it helpful to define a lending cut-off or threshold as a guideline for either more junior credit officers or for pre-screening in loan underwriting. Such institutions require a simple rule for defining a cut-off above which credit will be granted and below which it will be denied. Other institutions desire a rational pricing scheme for lending.

It can be difficult to derive information about economic value and subsequent usage policies directly from the statistics typically used to report default model performance. By convention, credit default models are evaluated using *power curves* which quantify the models' predictive power (c.f., Birdsall, 1966, 1973; Hanley and McNeil, 1982; Hanley, 1989; Pepe, 2002; Sobehart et al., 2000; Swets, 1988, 1996). A *receiver operator characteristic* (ROC) curve [1] plots the Type II error against one minus the Type I error. In the case of default prediction, it describes the percentage of non-defaulting firms that must be inadvertently denied credit (Type II) in order to avoid lending to a specific percentage of defaulting firms (1-Type I) when using a specific default model. Since there are (usually large) costs associated with extending credit to defaulting firms and (usually smaller) costs associated with not granting credit (or granting credit with overly restrictive terms) to subsequently non-defaulting firms, ROC analysis produces a form of cost–benefit analysis.

The optimal cut-off (i.e., the cut-off that minimizes costs) can be determined through standard ROC analysis. The approach permits the relative value of two models to be quantified and can be extended to accommodate real-world conventions such as relationship lending. This analysis can also be related to the lending practice of granting credit when a positive NPV is expected.

Although cut-offs are often used in practice, choosing a pass/fail type cut-off is almost always sub-optimal. As it turns out, however, ROC curves can also be used to derive basic pricing relationships. These are lower cost and more flexible than simple cut-offs. As a single decision lending criterion, a cut-off is a coarse measure when compared with probability based pricing schemes such as those we develop in the latter part of this article.

Although we do not focus on it in this article, it is important to recognize that all single-decision criteria are incomplete from a portfolio perspective as they ignore correlation and the impact of a particular credit within a portfolio context. [2] More generally, lending approaches can be divided into the following categories:

[1] In this paper we discuss ROC curves, but power curves are generally known by a variety of names and in a number of variations (e.g., cumulative accuracy profile (CAP) plots, etc.) These representations are typically isomorphic given appropriate information.

[2] For a discussion of these effects, and how they can be measured, see, for example, Kealhofer and Bohn (2001).

- cut-offs and prices based on stand alone risk measures;
- cut-offs and prices based on portfolio-based risk measures;
- market-based pricing; and
- portfolio-based pricing.

In this article, we focus on the first steps in this process as it is the norm for many institutions today.

The remainder of this article is organized as follows: In Section 2 we discuss some of the background of ROC curves and how they relate to profitability. In Section 3 we give an example of how cut-offs might be set in various lending environments. Section 4 shows how essentially the same mechanics that are used for defining cut-offs can be used to calculate appropriate prices and that this pricing corresponds to traditional NPV based lending approaches. Section 5 shows *why* more powerful models have more economic value, irrespective of whether an institution is using a simple lending cut-off or performing more complete probability based pricing calculations. Section 6 presents some simulation experiments that demonstrate *how much* economic value a difference in power can be worth. Section 7 gives a summary and conclusions.

2. ROC analysis and optimal cut-offs

Perhaps the most basic approach to understanding the performance of a default prediction model is to consider the number of predicted defaults (non-defaults) and compare this with the actual number of defaults (non-defaults) experienced. A common means of representing this is a simple *contingency table* or *confusion matrix*.

	Actual default	Actual non-default
Bad	TP	FP
Good	FN	TN

In the simplest case, the model produces only two ratings (bad/good). These are shown along with the actual outcomes (default/no default) in tabular form. The cells in the table indicate the percentage of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN), respectively. TPs and TNs are respectively defaults and non-defaults that are predicted correctly. FP is a predicted default that does not occur and a FN is a predicted non-default that actually defaults. The errors of the model are shown on the off diagonal where FN represents a Type I error and FP represents Type II error.

For models that produce more than two ratings or continuous outputs such as probabilities, a two-way contingency table can only be constructed for a specific model cut-off point. For example, a bank using a model that produces scores from one to ten might introduce a policy of underwriting loans to firms with model scores better than five. In this case, the TP cell would represent the percentage of defaulters

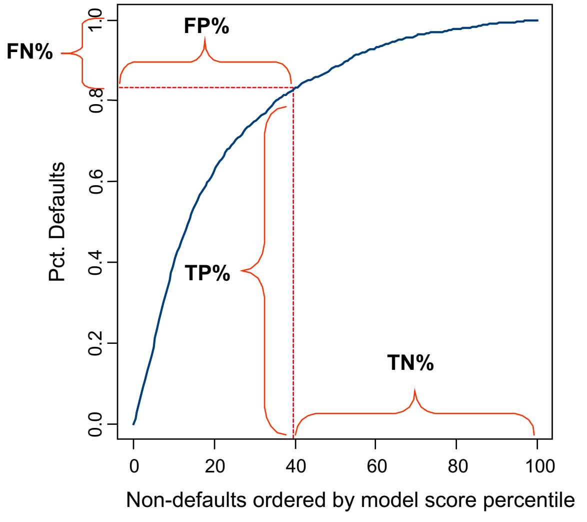

Fig. 1. Schematic of a ROC. This figure shows how all four quantities of a contingency table can be identified from a ROC curve. Each region of the x and y axes have an interpretation with respect to error and success rates for defaulting (FN and TP) and non-defaulting (FP and TN) firms. For example, to exclude 83% of the defaults using this model one would also have to forgo lending to 40% of the non-defaulters.

worse than five and the FP would represent the percentage of non-defaulters worse than five.

For some models, different cut-offs will imply different relative performances. A conservative cut-off might result in a contingency table that has lower errors for Model A while a liberal cut-off might result in a contingency table that favors Model B, and so on. Thus, using contingency tables (or indices derived from them [3]) to evaluate models can be challenging due to the relatively arbitrary nature of cut-off definition.

Fig. 1 provides an example of a ROC curve for a default prediction model. In the figure, all of the firms in the sample are scored with the model and the non-defaulting firms are ordered from worst (left) to best (right) along the x axis. The y axis shows the percentage of defaults avoided by setting a cut-off at a specific x value. The example indicates how the FN, TN, FP and TP rates can be calculated for a specific cut-off value. In the figure, the cut-off is the 40th percentile of the model's scores. Thus, the error rates correspond to a strategy of only lending to borrowers with scores better than the best 60% of non-defaulting firms, according to the model. Note that TP and TN always range from 0% to 100%. In other words, ROCs are invariant under assumptions about the baseline default rate, given a specific testing sample. The baseline default rate in a sample does not affect the shape of the ROC.

[3] For example, Swets (1996) lists ten of these summary measures and provides an analysis of their correspondence to each other.

Cut-off points may be based on operational business constraints (e.g., there is only enough staff to follow x% of the total universe), but these are typically sub-optimal from a profit maximization perspective. A more rigorous criterion can be derived with knowledge of the prior probabilities and the cost function, defined in terms of the costs (and benefits) of FN and FP (and TN and TP). [4]

The cost to the organization of a specific type of error (FP or FN) or benefit of a correct prediction (TP or TN) is independent of the model being used. However, the total payoff associated with using a *specific strategy* is dependent on both the cost to the organization of an error *and* the performance of the model. All things being equal, the higher the error rate of a model the more costly it is to make decisions that rely upon it.

The cost associated with using a particular model to make pass/fail lending decisions is given as the probability-weighted sum of the costs and benefits of the errors associated with the strategy. In our context, a strategy will be defined as a combination of a specific model and a specific cut-off (m,k). Thus the cost, C_s, of using a strategy, s, based on a combination of specific threshold (cut-off) k with a particular model m, is given as the probability-weighted sum of the costs and benefits associated with the rule defined by the cut-off, k:

$$C_s = p(D) \cdot c(\mathrm{FN}) \cdot \mathrm{FN}_{m,k} - p(D) \cdot b(\mathrm{TP}) \cdot \mathrm{TP}_{m,k} + p(\mathrm{ND}) \cdot c(\mathrm{FP}) \cdot \mathrm{FP}_{m,k} - p(\mathrm{ND}) \cdot b(\mathrm{TN}) \cdot \mathrm{TN}_{m,k} \quad (1)$$

where $b(\cdot)$ and $c(\cdot)$ are the benefit and cost functions, respectively, $p(\cdot)$ is the unconditional (population) probability of an event, [5] D and ND are default and non-default events, respectively. Thus the cost function is the expected benefit of correct decisions less the expected cost of mistakes. Translating into quantities of the ROC (see Fig. 1) we get:

[4] Although we take them as given in this discussion, identifying the drivers of these costs is typically non-trivial for banking institutions and is usually institution specific. In general, the costs of a FP are typically far lower than that of a FN. For example, Altman et al. (1977) conducted one of the first studies on this topic. The authors found that the ratio of the FN:FP costs was on the order of about 35:1. The appendix to this article gives an example of how a bank might begin to size these costs as well as how relationship pricing might enter the cost structure.

[5] In this context, it is important to have knowledge of the probability distribution of defaulters and non-defaulters in the population (the baseline default rate). This is because TP is defined as the percentage of all defaults captured at a particular score. TP thus ranges from 0 to 1. Consider the following case: If the default rate were extremely low (say 10 bps), then a model that predicts no-default in every instance would have an overall "accuracy" of 99.9%, where here accuracy means the percentage of misclassifications. However, that same model would have an accuracy of only 75% if the default rate were 25%. In both cases, since the model is the same, the ROC curve would be identical, but the observed performance of the model on the portfolio would be different as a result of the baseline default rate being different. If, as is usually the case, lending to defaulters were very much more costly than not lending to non-defaulters, a bank's tolerance for the model's error rate might be different depending on the baseline default rate.

$$\mathrm{FP}_{m,k} = k,$$
$$\mathrm{TN}_{m,k} = 1 - k,$$
$$\mathrm{FN}_{m,k} = 1 - \mathrm{ROC}(k),$$
$$\mathrm{TP}_{m,k} = \mathrm{ROC}(k),$$

$$C_s = [p(D) \cdot c(\mathrm{FN}) \cdot (1 - \mathrm{ROC}(k))] - [p(D) \cdot b(\mathrm{TP}) \cdot \mathrm{ROC}(k)] + [p(\mathrm{ND}) \cdot c(\mathrm{FP}) \cdot k] - [p(\mathrm{ND}) \cdot b(\mathrm{TN}) \cdot (1 - k)]. \tag{2}$$

Setting to zero, differentiating C_s with respect to k and rearranging terms, gives the slope of a line with marginal cost equal to zero.

$$S = \frac{\mathrm{dROC}(k)}{\mathrm{d}k} = \frac{p(\mathrm{ND})}{p(D)} \frac{[c(\mathrm{FP}) + b(\mathrm{TN})]}{[c(\mathrm{FN}) + b(\mathrm{TP})]}. \tag{3}$$

The point at which the line with slope S, defined above, forms a tangent [6] to the ROC curve for a model will define the optimal cut-off, given a particular set of costs and benefits, as this point will be the one at which marginal payoffs (costs) are zero. Green and Swets (1966) provides a discussion of this approach (chapter 1) and an analytic formulation of the problem as applied to ROC analysis. The authors show that for any ROC curve and cost function, there exists a point with minimal cost at which both the Type I and II errors are minimized within the constraints of the cost function. [7]

A central assumption is that there is a correspondence between the model scores and default probabilities. Fortunately, there is a direct relationship between calibration and power. [8]

A line with a slope defined as in (3) has been termed an *iso-performance* line (Provost and Fawcett, 1997). We can use a ROC curve to determine where the optimal cut-off score will be for a given cost function. It is the point where an iso-perform-

[6] We use the term tangent somewhat informally in this context since in practical settings the "curve" defined by a ROC may be piecewise linear or otherwise discrete, and thus be non-differentiable. The term tangent is used to convey the conventional meaning when interpreting ROC curves as if they were continuous and once differentiable. If the ROC has discontinuities or is otherwise non-differentiable, numerical approaches can be used. (See Fig. 8 for a simple example.)

[7] For a more general discussion of decision-based methods to model performance evaluation, see Granger and Pesaran (1999).

[8] While it is not always the case that powerful models are calibrated accurately to probabilities of default, it is empirically feasible to calibrate a model to real-world default probabilities by performing a default study on the historical behavior of each model score. Thus assuming a bank has historical data and can perform ROC analysis, it can also calibrate a model to a similar level of accuracy using the same machinery. In general and if calibrated correctly, more powerful models produce more accurate probability estimates as well (cf. Stein, 2002).

ance line with slope S forms a tangent to the ROC curve for a particular model. This is the point, k^*, that will minimize this cost function for the particular model. [9]

This methodology is a straightforward application of cost minimization (setting marginal costs equal to zero). However, since this perspective is not always obvious in the context of ROC analysis, [10] it can be useful to frame the discussion in terms of more traditional lending practices.

Consider how a lending officer typically views a lending decision in NPV terms. Assuming sufficient capital, the officer will make all loans where the NPV of the expected cashflows is positive. If the values of the payoffs are assumed to be given in NPV terms, and the unconditional probability of default is π, then the NPV without a model would be evaluated as

$$\mathrm{NPV} = (1-\pi)V_{\mathrm{ND}} + \pi V_{\mathrm{D}}$$

where V_{D} and V_{ND} are the values of the payoff in the event of default and non-default respectively.

Substituting

$$\begin{aligned} &\pi = p(D),\\ &(1-\pi) = p(\mathrm{ND}),\\ &V_{\mathrm{ND}} = b(\mathrm{TN}) - c(\mathrm{FP}),\\ &V_{\mathrm{D}} = b(\mathrm{TP}) - c(\mathrm{FN}) \end{aligned}$$

we get:

$$\mathrm{NPV} = p(\mathrm{ND})\cdot b(\mathrm{TN}) - p(\mathrm{ND})\cdot c(\mathrm{FP}) + p(D)\cdot b(\mathrm{TP}) - p(D)\cdot c(\mathrm{FN}).$$

How does this change if we introduce a credit scoring model? The model provides additional information about the conditional (borrower-specific) probability of default and allows the user to define these probabilities explicitly. To evaluate the NPV of a loan to a firm that scored in the kth percentile of all firms using model m we can use:

[9] Again, if the ROC curves for two models cross, then neither is unambiguously "better" than the other with respect to a general cut-off. The preferred model will depend critically on the cost function. When one ROC completely dominates the other (i.e., the ROC curve for the dominant model is above the ROC for the other model), in contrast, the dominant model will be preferred for any possible cut-off chosen. For some types of applications, it may be possible to use the two models to achieve higher power than either might independently through the creation of a ROC convex hull that covers both models (Provost and Fawcett, 1997), although such approaches may not be suitable for real-world credit problems since they require a probabilistic (non-deterministic) choice to be made between the two models for each borrower. However, if this strategy is implemented in the case where the ROCs of two models cross, the expected performance of the combined two models will be superior to either used separately.

[10] For example Baestaens (1999) discusses a number of useful approaches to implementing internal rating grades based on commercial models. The objective of the article was to optimize performance over several criteria. The article also proposes minimizing Type I error exclusively in determining a "red flag" cut-off. As we show here, except in the special case where this criterion coincides with the minimum of the true cost (of Type I and II error) function, this will be sub-optimal. In the context of a multi-objective optimization problem, this is less clear.

$$\text{NPV} = p(\text{ND}) \cdot b(\text{TN}) \cdot \text{TN}_{m,k} - p(\text{ND}) \cdot c(\text{FP}) \cdot \text{FP}_{m,k} + p(D) \cdot b(\text{TP}) \cdot \text{TP}_{m,k} - p(D) \cdot c(\text{FN}) \cdot \text{FN}_{m,k}$$

and we are left with the negative of Eq. (1); negative costs are positive cashflows. Thus, assuming a constant cost function and setting a cut-off at the point where the marginal NPV is zero is equivalent to setting a cut-off at the point on the ROC where a line with slope S forms a tangent. A bank should continue to make loans until the marginal return for doing so is zero.

3. A short example

In this example we demonstrate how iso-performance lines are used to determine the percentile k^* of an optimal lending cut-off score and examine the expected cost of strategies based on k^*. We assume two models of default, Models P and W, where P is more powerful and W is weaker. For convenience, we assume that the model scores have been sorted and that the score $s_{(k)}$ is the value of the model output in the kth percentile of all non-defaulters. (Thus if 25% of all non-defaulting firms had a score lower than 3 then $k = 25$, and $s_{(25\%)} = 3$.)

In the remainder of this article we assume for tractability that the lending bank grants loans that mature in one year, making one payment at the maturity date and that defaulting firms default at maturity without paying accrued interest. We do this to reduce the complexity of the calculations of NPV as this eliminates all probabilistic components of the costs, with the exception of the default rates.

Unless noted otherwise, we use the assumptions in Table 1 as the baseline case in our examples. These assumptions simplify somewhat the complexities of many lend-

Table 1
Assumptions for baseline underwriting costs and profits

	Variable	Baseline value
1	$p(D)$	2.0%
2	$p(\text{ND})$	98.0%
3	Interest spread (per annum)	1.25%
4	Underwriting fees (up front)	0.50%
5	Workout fees (on default)	2.0%
6	LGD (on default)	35.0%
7	Risk free rate	4.00%
8	Additional relationship benefit (NPV)	0

The figures in this table show the (simplified) assumed costs (and profits) for underwriting to a "typical" client of a bank. Rows 1–2 indicate the assumed baseline probability of default (and its complement) in the bank's market. 3–4 represent the fees and revenue the bank will generate by making a loan and 5–6 the costs associated with a default. Note also that in this case there is no relationship value to making or not-making a particular loan, but that if there were, it would be captured in NPV terms for the life of the loan on line 8. All costs and revenues are quoted as percentages of a dollar loaned.

ing situations, yet they provide enough detail for a realistic perspective on the approach we describe.

3.1. The case of uniform terms for all clients

We begin by examining the case in which the lender does not distinguish between "preferred" clients (i.e., those that do a high volume of business with the bank) and other clients. In this case, a single cost function is applied to *all* clients and the k^* for a particular model is also the same for all clients.

Using the base case assumptions for costs generates the iso-performance lines shown in Fig. 2. In this case, k^* is equal to the value of the model that is associated with the percentile of the distribution defined by the x value of each point. For example, here, the x value associated with the tangent for the weaker model is the 12th percentile. Thus here $k^* = 12\%$ and $s_{(k^*)} = s_{(12\%)}$. In this case, we examined a model that produces scores on a scale from zero to twenty. For this model, it turned out that $s_{(12)} = 14$. Thus, a bank with the cost function described here using this model (Model W) would set its cut-off at 14 and would grant credit to those borrowers with scores lower than 14 and deny credit to those with higher scores.

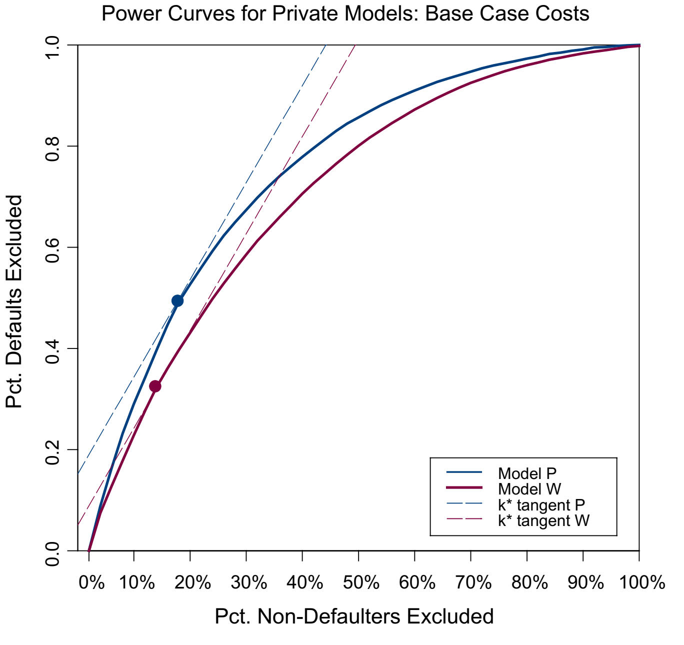

Fig. 2. Iso-performance lines that define k^* under base case assumptions. This figure shows two models and the iso-performance lines that define the optimal cut-off for each under the assumptions outlined for the base case. Dots indicate the point of tangency (and x values) that define k^*.

The analysis shows that the difference in performance of applying k^* in each case can be substantial. In this example, the more powerful model, Model P, would exclude 45% of the defaults at an approximate profit of 104 basis points (per dollar approved). Model W makes a bit less (99 bps) and only excludes about 28% of the defaults (as shown by the y positions of the tangent points on the ROC).

In the next example, we use the base case assumptions for costs but assume the default rate in the bank's market is 5% rather than 2%. As before, we calculate the iso-performance lines, shown in Fig. 3, which define the optimal cut-offs. Note how much flatter the lines are, as a result of the higher cost associated with being wrong in predicting defaults. These result in k^* values that are much more conservative than in the base case. For example, the cut-off percentile for Model P is now the 39th percentile versus the 8th in the base case example.

Here again, the analysis shows the difference in performance of applying k^* in each case to be substantial. In the second example, the more powerful model and the weaker model exclude about 80% and 78% of the defaults, respectively, but the costs are very different. In the case of the powerful model, the approximate profit is 62 bps. This turns out to be about 24% higher than in the case of the weaker model which has a profit of around 50 bps. Here again, not only does the weaker model exclude fewer of the defaults, but it does so at a higher cost.

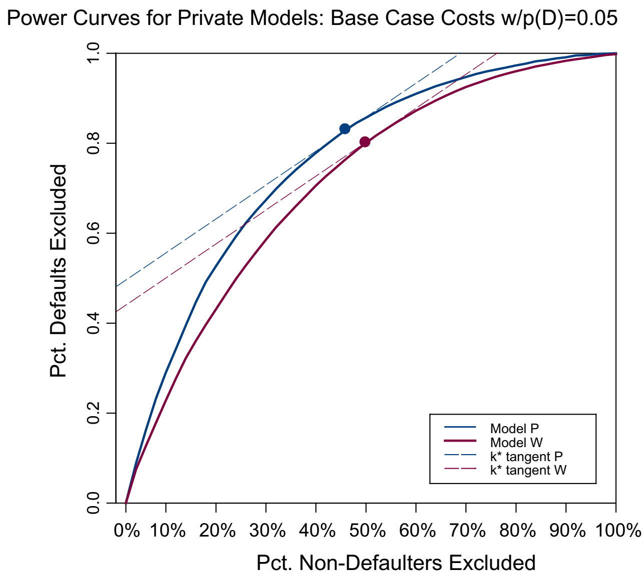

Fig. 3. Higher default rate case tangential iso-performance lines that define k^*. This figure shows two models and the iso-performance lines that define the optimal cut-off for each under the assumptions outlined for the base case except that the default rate has been increased to 5%. Dots indicate the point of tangency (and x values) that define k.

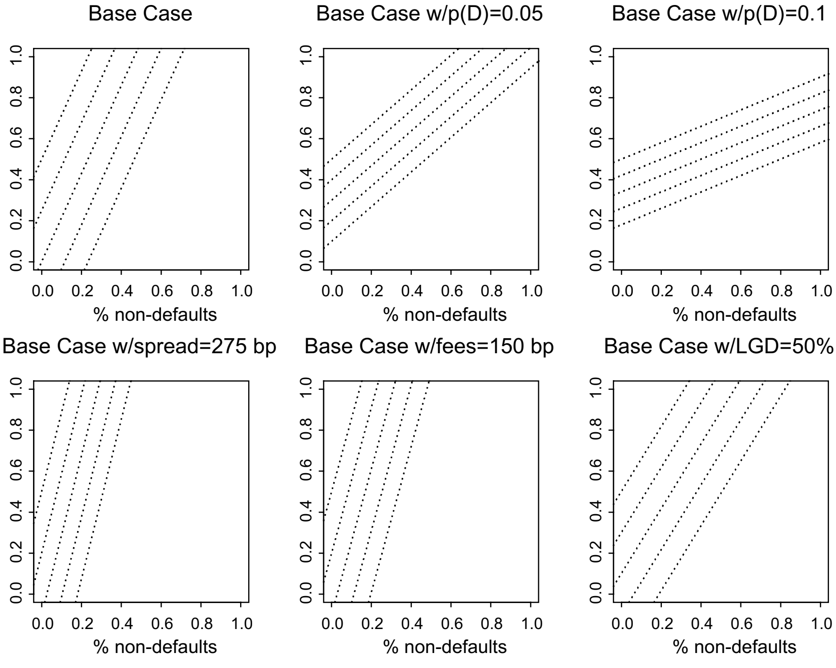

Fig. 4. Iso-*performance* lines for variations on the baseline cost assumptions for loan underwriting. These figures show how the slope of iso-performance lines change as various assumptions are modified. The base case from Table 1 is shown in the upper leftmost panel. The remaining cases change various of the assumptions in the base case. Note that increasing the "loss" component of the cost function (higher default rate, higher LGD, etc.) decreases the slope of the iso-performance line and that increasing the "opportunity cost" component of the cost function (higher underwriting fees, wider spreads, etc.) increases the slope.

To better understand how iso-performance lines behave, consider the examples given in Fig. 4 which show how the slopes change as we vary certain assumptions about the costs of underwriting loans to potential borrowers. The panels of the figure show iso-performance lines for variations on baseline assumptions in Table 1.

It is obvious that increasing the "loss" component of the cost function (higher default rate, higher LGD, etc.) decreases the slope of the iso-performance line and that increasing the "opportunity cost" component of the cost function (higher underwriting fees, wider spreads, etc.) increases the slope. It is also clear that the steeper the slope the lower the cut-off k^* will be since it will "hit" the ROC further to the left on the ROC.

In general, intuition suggests that the higher the loss potential of a missed default the less steep the line and thus the more conservative the optimal cut-off becomes, defining strategies that grant credit to high scoring borrowers only. Similarly, the higher the opportunity costs in not granting credit to a non-defaulter, the steeper the iso-performance line becomes and thus the more lenient.

In Section 3.2, we extend the analysis to examine the case in which we have two classes of applicants, those who give the bank a high volume of adjacent business that could be lost if a loan is not granted and those who are not high volume clients.

3.2. Multi-tier cut-offs: Extending the framework to include relationship value

Consider the case in which a bank has a number of high-profile clients who are identical in all respects [11] to the body of the borrower population except that they provide the bank with significant revenue through the use of additional banking services. [12] What effect should this additional revenue have on the choice of k^*?

Fig. 5 shows this case under the assumption that the relationship is worth 50 bps per annum. In the figure, we show the iso-performance lines that define the optimal cut-off for both standard clients (dotted line) and relationship clients (dashed line). Note how k^* is less conservative for clients who also provide relationship revenue. For such clients the cut-off is around the 6th percentile whereas those clients without relationships are optimally cut-off at the more conservative 16th percentile.

It is not the case that the relationship clients get preferential treatment as a gesture of good will. Rather, since they provide more revenue when they do not default, through their other fees in addition to the loan related revenues, the opportunity cost of not lending to these borrowers is far greater than for the typical client. Thus, this analysis provides some justification for the practice of some bankers who relax lending practices for their better clients, but characterizes the additional cost of doing so as specific minimum additional fees required from adjacent business.

3.3. Zero cost lending cut-offs

Imagine now that the bank would like to lend at zero expected cost (or better), given a specific cut-off that it has chosen. For example, a bank may determine that it only has staff available to process a specific percentage of the potential applicants. In this case, the bank should optimize its lending policy for the specific cut-off value. If the bank is able to adjust the loan terms to reflect the risk of the specific cut-off policy, it can achieve more efficient lending by designing terms that result in a cost function that ensures that $k = k^*$, the optimal cut-off.

To determine the minimum profit that should be required for a policy based on setting a lending cut-off to a specific value of k, we need to ensure that the marginal revenues at each risk level fully compensate the bank for expected losses. We then solve for $b_{k^*=k}(\text{TN})$, the amount of benefit (spread + underwriting fees in this

[11] If this were not the case, the same ROC would not be appropriate for both classes.

[12] Note that this additional value can in some cases be viewed as a real option. Consideration of such real options can lead to significantly different lending decisions at the borrower level than those that would be taken in absence of such options.

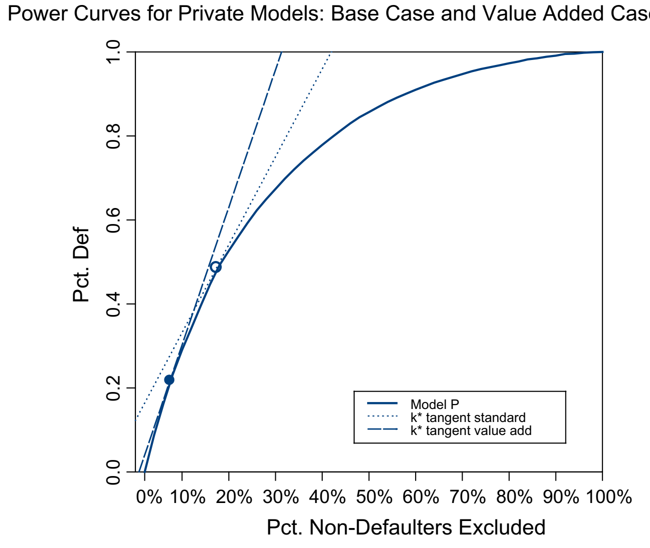

Fig. 5. Iso-performance lines that define k^* when clients provide relationship revenue. This figure shows the iso-performance lines that define the optimal cut-off for both standard clients (dotted line) and relationship clients (dashed line). Dots indicate the point of tangency (and x values); filled dot represents k^* for the relationship clients and the hollow dot represents k^* for standard clients. Note how k^* is less conservative in the case of clients also providing relationship revenue.

case [13]) required at each level of k in order to ensure zero costs (on average) for lending *if that point were used as a cut-off*.

Setting the overall cost, C, to zero and assuming $b(\text{TP})$ is also zero (no benefit for defaulters not granted credit), and rearranging, we get the following result for $b_{k^*=k}(\text{TN})$:

$$b_{k^*=k}(\text{TN}) = \frac{1}{p(\text{ND}) \cdot (1 - \text{FP})} [p(\text{ND}) \cdot c(\text{FP}) \cdot \text{FP}_{k,m} + p(D) \cdot c(\text{FN}) \cdot \text{FN}_{k,m}]. \tag{4}$$

This function produces the floor on the NPV of the total expected revenues that must be received to compensate for the associated risks of the lending environment described by the cost function. Importantly, this is the value of $b_{k^*=k}(\text{TN})$ that would be required if k were the cut-off, not if lending were done at all values of k (Fig. 6).

Put more clearly, $b_{k^*=k}(\text{TN})$ is *not the price* that should be charged for a loan with score $s_{(k)}$. Rather, it is the level of revenue that must be received (in NPV terms)

[13] Note that this is just a convenient adjustment as the lender can choose to divide the benefits among any sources desired.

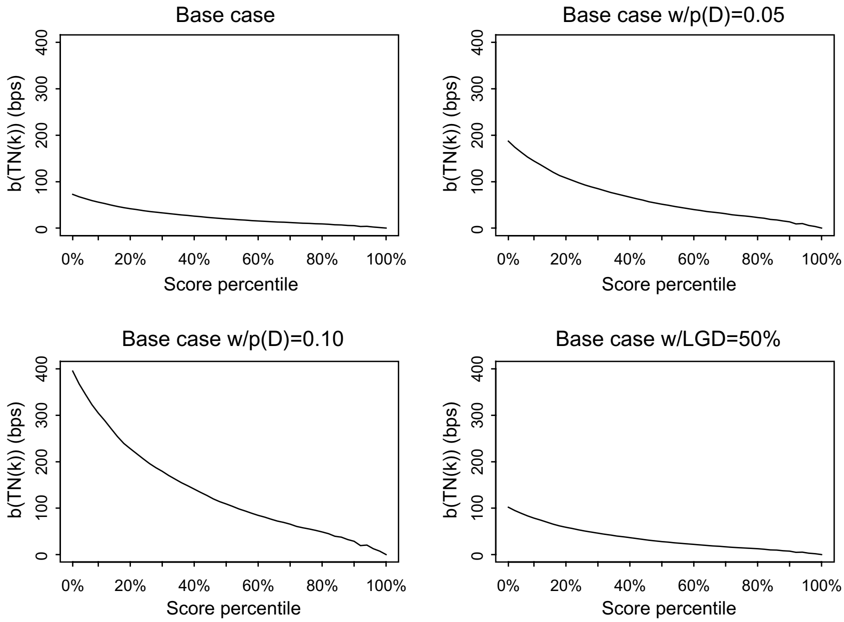

Fig. 6. Minimum required benefit to realize zero expected cost in lending for a fixed level k. This figure shows the minimum amount of benefit (e.g., spread, underwriting fees, etc.) that would be required to lend at zero expected cost at each cut-off level.

for *every borrower granted* assuming all borrowers with scores above $s_{(k)}$ are granted credit with uniform pricing and none are granted credit with scores below $s_{(k)}$. In other words, $b_{k^*=k}(\text{TN})$ is the value of $b(\text{TN})$ that would be required in a lending policy to make $k = k^*$, the optimal cut-off for that policy.

We note in passing that the approach described in this paper can be extended to a number of other policy applications. For example, in the case where an institution is, in fact, resource constrained and only has enough staff to review x% ($k = 1 - x$) of the applications, this analysis can also be adapted to determine the cost of setting $k = 1 - x$ versus allowing k to equal k^*. As another example, [14] the method could be used to set and evaluate interest rate margins between borrowing and lending.

4. A basic pricing scheme using ROCs and cost functions

We have shown how an optimal cut-off can be set for a particular cost function and, conversely, how a cost function can be designed to ensure that an arbitrarily

[14] This application was suggested by an anonymous reviewer.

chosen cut-off will be optimal. However, in both cases, the (restrictive) assumption was that all borrowers were given the same terms once the pass decision had been made.

In most practical settings, it is more reasonable to recognize the existence of heterogeneity in credit quality in the borrower population. In this case, the one-size-fits-all loan terms assumption needs to be relaxed so that each borrower can be given terms that appropriately compensate the bank for the credit risk of the borrower without unduly penalizing the borrower. In this section, we explore how ROC curves can be used to achieve differential pricing.

Recall now the relationship between ROC quantities and positive NPV lending:

$$\mathrm{NPV} = p(\mathrm{ND}) \cdot b(\mathrm{TN}) \cdot \mathrm{TN}_{m,k} - p(\mathrm{ND}) \cdot c(\mathrm{FP}) \cdot \mathrm{FP}_{m,k} + p(D) \cdot b(\mathrm{TP}) \cdot \mathrm{TP}_{m,k} - p(D) \cdot c(\mathrm{FN}) \cdot \mathrm{FN}_{m,k}.$$

Since in what follows we are concerned with pricing, we can simplify the above expression. In the case of a pricing approach, *no loans* are denied. Rather, any loan can be granted, provided the appropriate revenue (for which we must solve) is received for the loan with $s = s_{(k)}$. If no firms are turned down, lenders are only concerned with true negatives and false negatives. In order to determine the payoff for lending to a specific borrower with a specific score of $s_{(k)}$, we introduce some additional notation, letting tn and fn represent the derivatives of TN and FN with respect to k.

Setting the NPV to 0 and differentiating, we get

$$\frac{\mathrm{d}NPV}{\mathrm{d}k} = p(\mathrm{ND}) \cdot b(\mathrm{TN}) \cdot \mathrm{tn}(k) - p(D) \cdot c(\mathrm{FN}) \cdot \mathrm{fn}(k) = 0.$$

Rearranging yields:

$$b_k(\mathrm{TN}) \cdot p(\mathrm{ND}) \cdot \mathrm{tn}(k) = c(\mathrm{FN}) \cdot p(D) \cdot \mathrm{fn}(k)$$

where we now introduce a subscript on $b(\cdot)$ to indicate that, while costs of lending and default remain fixed, the benefits (interest spread, underwriting fees, etc.) must be adjusted to reflect the differing levels of risk associated with each $s_{(k)}$ since here we are charging a different price for each loan. This relationship is consistent with our general intuition: For correctly priced loans, the (probability weighted) benefit of the no-default case must be set equal to the cost of a default. (A proof of the relationship to NPV is given in Appendix B.)

Rearranging once again, we can solve for $b_k(\mathrm{TN})$, the total revenue that a loan must generate to break even, given its probability of default.

$$b_k(\mathrm{TN}) = \frac{c(\mathrm{FN}) \cdot p(D) \cdot \mathrm{fn}(k)}{p(\mathrm{ND}) \cdot \mathrm{tn}(k)} \tag{5}$$

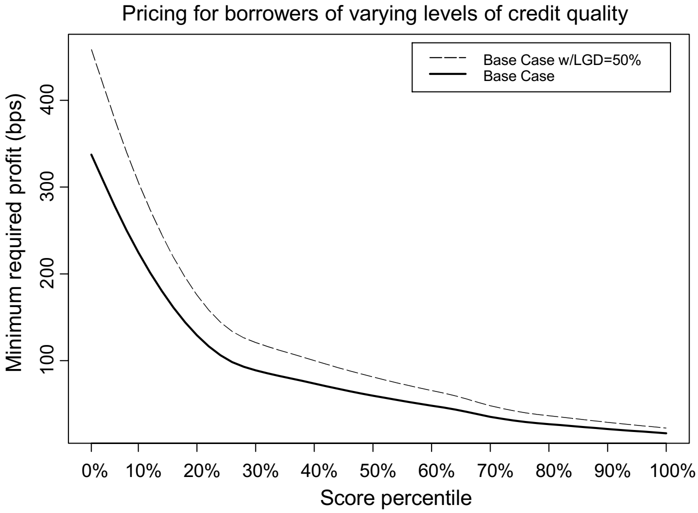

Fig. 7. Required profit at each value of k assuming lending to any borrower at zero cost. This figure shows the required minimum profit for loans with score $s_{(k)}$ using the powerful model. Note how the required profit shifts as expected losses increase (LGD increases).

Thus using only the ROC, information about the baseline default rates, and knowledge of the costs of lending, [15] we can set minimum acceptable prices [16] for lending. Note that we have moved from the ROC (cumulative) true positive and false negative to the density equivalent as we wish to evaluate the cost of lending to clients with probabilities of default that are equal to (rather than less than) the particular point on the curve. This permits tailoring the terms of the transaction to the exact risk of the borrower, rather than taking a single uniform set of terms for all borrowers (Fig. 7).

Also note that using this basic pricing mechanism, a bank would be able to grant credit to any borrower by adjusting the terms of the loan. (Lending regulations

[15] Realistic lending environments, loan terms, etc. would make the definitions of $b(\cdot)$ and $c(\cdot)$ considerably more complicated than we have shown in our examples.

[16] Note that we are using the term "pricing" somewhat loosely here. In particular, the probabilities on which the "prices" are based are not risk-neutral probabilities but rather actuarial or "real world" probabilities. That said, a more complete translation from real-world to risk neutral probabilities can often be made with an estimate of the risk premium associated with a particular borrower (cf. Duffie and Singleton, 2003). For example, a common approach to making this transformation for a one period probability of default involves the following relationship:

$$\pi^* = \Phi(\Phi^{-1}(\pi) - \eta)$$

where π is the actuarial probability of default, π^* is the risk neutral probability of default, Φ and Φ^{-1} are the cumulative and inverse cumulative standard normal distribution functions respectively, and η is the risk premium for the borrower. (For longer periods this scales with the square root of time.) Also note that in some presentations η enters through addition rather than subtraction.

might prohibit the bank from realizing revenues in some cases, though, due to restrictions on the permissible levels of interest rates, etc.)

Here we have shown that pricing curves can be derived from a fixed cost function and an ROC, but it is straightforward to extend the approach to arbitrary cost functions that need be neither linear nor constant. If, rather than defining the pricing curve, we instead take only $p(D)$, $\text{fn}(k)$ and $\text{tn}(k)$, we can use these directly in a cost function to derive a price.

For example, it is reasonable to contemplate a borrower-specific (non-constant) LGD. In this case, the price for a loan to the ith borrower, with score $s_{(k)}$ might be given as

$$b_k(\text{TN}) = \frac{c_i(\text{FN}) \cdot p(D) \cdot \text{fn}(k)}{p(\text{ND}) \cdot \text{tn}(k)}$$

where $c_i(\text{FN})$ is now the *borrower specific* cost of default.

5. More powerful models are more valuable

When using a fixed lending cut-off, lenders assign a borrower to one of two categories based on whether its score is above or below a threshold. Fig. 8 shows the cost function for several of the example scenarios, evaluated at every value of k. Using the

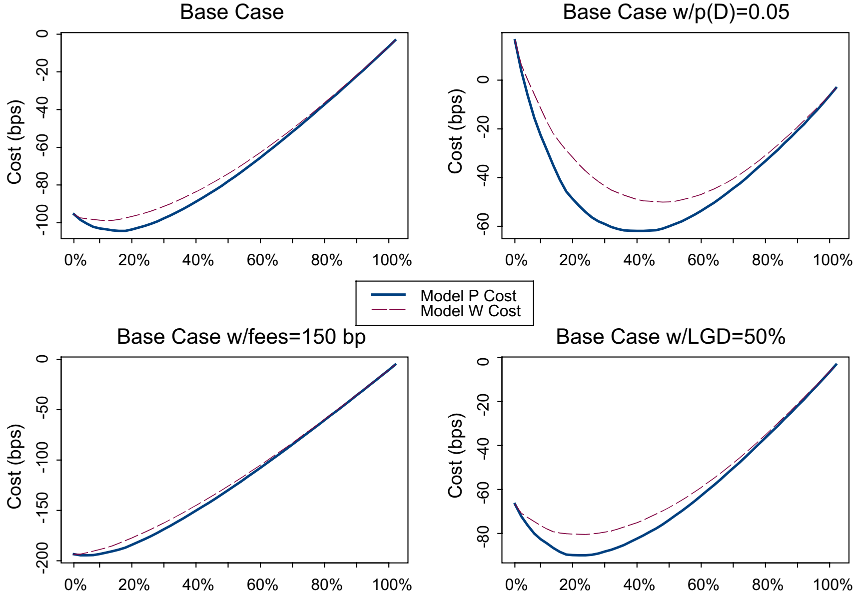

Fig. 8. Cost function for different assumptions evaluated at various levels of k. This figure shows the estimated cost functions defined by the scenarios discussed above. The functions are evaluated at all values of k. It is clear that for any value of k, the more powerful model has a higher payoff (more negative cost) than the weaker model.

weaker and more powerful models from our examples, it is clear that for any value of k the (uniformly) more powerful model has a higher payoff than the weaker model. We would expect this result, because when the ROC curves do not cross, the more powerful model will always have lower Type I error for the same Type II error, resulting in a higher payoff, assuming that costs and benefits are both positive.

A similar result obtains in the case of pricing policies based on default prediction models. Users of a weak model will systematically over (under) charge higher (lower) credit quality borrowers, creating competitive opportunities for those who use the more powerful models. This can be detrimental in the case of relationship lending, since good customers may be asked to pay more than they should using pricing policies based on a weaker model. Conversely, in the riskiest portions of the portfolio the lender using a weaker model will often not be adequately compensated for the levels of lending risk.

Importantly, since models with different power, by definition, rank order borrowers differently, it is not possible to infer from this analysis what the differential in pricing would be for a *particular* borrower without actually evaluating the borrower. Even if both banks were using ROC analysis to set prices, for example, it is possible that each bank might give the borrower different relative scores. While Bank A using model A might classify a borrower as in the riskiest 20% of the population, Bank B might classify the same borrower as being in the riskiest 10% using Model B. Pricing differentials for the particular borrower would arise then as a result of both the differential in pricing curves between the models *and* (possibly) the different rankings of firms by the models.

6. Some simulation results

A companion paper to this one (Stein and Jordão, 2003), presents a methodology for extending the approach described in this paper to a simulation framework that generates a distribution of expected values for the use of different models. The methodology provides a means of estimating the typical benefit (cost) in cash terms of using a more powerful model over a weaker one. We discuss some of the particulars of this approach in more detail below.

To give a flavor of the simulation results here, we show the performance of two models that differed in accuracy ratio [17] by about 5 points. We note by way of background that the more powerful model tested here was a financial statement-based statistical model developed using an approach similar to that used in a class of models known as *generalized additive models* (Hastie and Tibshirani, 1990). The second model was the four variable [18] Z-score model of Altman (see Altman, 2000, for a discussion).

[17] Accuracy ratios are summary statistics for CAP plots (Sobehart et al., 2000) and are isomorphic to the area under the ROC.

[18] Note that while this model was developed for non-manufacturing firms, our testing showed it to be more uniformly powerful at predicting default for middle-market firms than other variants of the Z-score models, regardless of their industry.

We drew data from Moody's Credit Research Database (CRD), a database of middle-market borrower and loan performance information collected from banks and other institutions worldwide. Importantly, these borrowers represent firms that actually applied for bank loans from commercial banks. To estimate the impact of differences in model performance, we sized the lending volume of a typical mid-sized US bank (about \$50B in consolidated assets) based on data from the CRD. In our data set, a typical bank underwrote about \$4.25B in loans (new and renewal) in 2002 and we took this as a proxy for average annual lending activity in the middle-market.

In interpreting the simulation results, note that the specific outcomes we describe here do not generalize beyond the two models, economic conditions and cost structures we examined. In other words, it is not possible to draw specific inferences about the value of a particular power differential (e.g., 5 points of accuracy ratio) and a particular economic benefit since the value of one model versus another is sensitive to both the economic conditions under which the models would be used, and the shapes of the two power curves.

For example, if two curves cross but have the same power (accuracy ratio) the placement of a cut-off may favor one or the other model, depending on which side of the crossover point it is placed. Individual models need to be tested directly to determine the economic impact under specific cost structures.

As a first experiment, we tested the impact on a bank's profits of simply switching from a weaker to a more powerful model. To do this, we constructed a large number of sets of loan applicants by randomly choosing sets of firms from the North American CRD.

We scored each of the random sets of applicants first with the weaker model and then with the more powerful model. We then compared the performance of the portfolio the bank would have gotten using the weaker model to that of the more powerful model using an optimal (under each model) cut-off policy. Note that for any set of applicants, the bank's profitability varies only due to the choice of the model and cut-off: Differences in the portfolio performance under one policy or the other are due to the difference in the set of applicants that would have been granted credit or denied credit. In the simulation the mean difference in weak portfolio and powerful portfolio performance was about 4.1 basis points. In this case, we can check the consistency of the simulated result with the analytic result. We calculated the expected difference in cost in analytic form using (1) and setting $k = k^*$ for each model. From (1) we estimated that a bank using simple cut-offs would save approximately 4.5 basis points (per dollar originated) if it switched to the more powerful model from the weaker one. The simulated performance using the empirical data (4.1 bps) was reasonably close to our analytic estimate of 4.5 basis points which was well within the sampling variation ($\sigma = 0.9$ bps) of the simulation.

Based on these simulation results, we estimated that the impact of the switch to the more powerful model would be about \$1.7MM per year assuming the lending pattern of the prototypical bank we described.

Because banks compete for deal flow, a more realistic test of both the cut-off and the pricing approaches is to consider what a rational borrower would do if it were to seek credit. Since models of different power cannot be not perfectly correlated, a borrower that might be rejected by one bank might be accepted by another if the banks were using different models. This is true even in the case where both banks were using optimal (for their model) cut-offs. Similarly, the risk category assigned to a borrower for pricing by one model might be different than the risk category assigned to it using another model and this could lead to differential pricing between the two banks. All things being equal, a rational borrower would opt for the bank offering cheaper loan terms.

Stein and Jordão (2003) extends the simulation approach to examine how this more complex behavioral relationship relates model power to economic benefits. The details of the methodology are beyond the scope of this article, but the results are informative. For instance, the results of that study indicate that differentials in performance are more pronounced in a competitive setting than in a simple switching setting.

As an example, the simulations showed that if the bank in the above example switched to the more powerful model but a competing bank did not, the revenue differential would be about 11 bps, considerably higher than the 4.1 bps expected when the bank simply switched in isolation. Furthermore, if the banks were competing using a pricing approach, rather than the cut-off approach, the difference in return would be on the order of about 16 bps, which would equate to approximately $6.8MM per year for a prototypical medium sized bank.

More generally, the experiments showed a very conservative estimate of the additional profit generated per point difference of accuracy ratio to be 0.97 bps when using the cut-off approach and 2.25 bps when using the pricing approach. Thus, a conservative estimate of the additional profit that a bank could expect using a model five points of accuracy ratio better than its competition would be around five basis points per dollar granted, if the bank were using the cut-off approach, and eleven basis points per dollar granted under the pricing approach. For the prototypical medium sized bank described earlier, this would equate to additional profits of about $2.1 million and $4.8 million, respectively, in 2002.

7. Conclusion

We have provided some intuition for why more powerful models can set prices more accurately than weaker ones. Lenders using more powerful models will have an advantage over those that do not as they are able to align their fees and loan terms more exactly and thus avoid over- or under-charging clients.

The framework we presented can be used to establish the optimal cut-off point for lending decisions based on a lender's cost function. This cut-off point minimizes costs relative to all other possible cut-off points. This criterion results, *ceteris paribus*, in more conservative thresholds when default risk is higher or opportunity costs associated with foregoing lending revenue are lower. The approach can be used to reflect realistic details of the lending such as workout and recovery costs, foregone profits on loans that would not have defaulted, and relationship lending.

This basic cut-off approach can be easily extended to a basic pricing scheme that is more flexible. Such a pricing scheme relates the statistical concepts of ROC analysis directly to pricing in NPV terms with which many bankers are familiar. We have also shown how non-constant cost functions might be accommodated to allow for borrower-specific LGD, etc.

A consequence of this analysis is that *powerful models cost less to operate than weaker ones*, given a fixed cost function. As a corollary, organizations that use more powerful models can do more profitable lending (even in the case where fewer loans may be granted) than organizations that use weaker models.

Although we do not recommend cut-off setting, preferring instead more complete probabilistic pricing approaches, it is still widely practiced among lending institutions. We have provided some guidance on how to set such thresholds if lenders determine that they must. We have also shown how the same machinery that is used to determine optimal cut-offs can be extended to produce prices that permit lending to any borrower, given appropriate expected revenue generation. This is useful since even k^*, the optimal cut-off value for pass/fail decisioning, is only optimal within the context of this decision process. The cut-off-setting process itself still results in a sub-optimal lending policy.

A more complete pricing approach would also incorporate the correlation of a new loan with the existing portfolio into a pricing model that also reflected the overall market price of risk.

Acknowledgements

I am grateful to David Lando of the University of Copenhagen, Foster Provost of NYU and Alan White of the University of Toronto, to Jeff Bohn, Jon DiGiambattista, Greg Gupton, Jim Herrity, Ahmet Kocagil and Bill Morokoff of Moody's KMV, to Richard Cantor, Lea V. Carty of Moody's Investors Service, and to two anonymous referees for their generous comments. I am especially indebted to Chris Mann of Moody's Investors Service and to Douglas Dwyer of Moody's KMV who provided important insights into aspects of the pricing formulation. All remaining errors are of course my own.

Appendix A. Calculation of iso-performance line in base case scenario

The following appendix provides some detail of the iso-performance line calculation for the base-case scenario. While we have made certain assumptions about how to classify various cashflows, other representations of $b(\cdot)$ and $c(\cdot)$ could easily have been used. The purpose here is to demonstrate the basic approach, rather than recommend a particular formulation as appropriate for a specific situation.

We have reproduced Table 1 for convenience and show the calculation that we have chosen for the various quantities $b(\cdot)$ and $c(\cdot)$.

Assumed values of costs and benefits and calculation of S:

$$
\begin{aligned}
b(\mathrm{TN}) &= (4) + (3) + (8),\\
b(\mathrm{TP}) &= 0,\\
c(\mathrm{FN}) &= \mathrm{NPV}[(5) + (6)] + (8) - (4)\\
&= [(5) + (6)]/[1 + (7)] + (8) - (4),\\
c(\mathrm{FP}) &= 0,\\
b(\mathrm{TN}) &= 1.75\%,\\
b(\mathrm{TP}) &= 0.00\%,\\
c(\mathrm{FN}) &= 35.08\%,\\
c(\mathrm{FP}) &= 0.00\%,
\end{aligned}
$$

$$
\begin{aligned}
S &= \frac{p(\mathrm{ND})}{p(D)} \frac{[c(\mathrm{FP}) + b(\mathrm{TN})]}{[c(\mathrm{FN}) + b(\mathrm{TP})]},\\
S &= \frac{0.98}{0.02} \frac{[0 + 0.0175]}{[0.3508 + 0]},\\
S &= 2.44.
\end{aligned}
$$

Appendix B. The relationship between NPV terms and ROC terms

We wish to show that

$$b_k(\mathrm{TN})p(\mathrm{ND})\mathrm{tn}(k) = c(\mathrm{FN})p(D)\mathrm{fn}(k) \tag{B.1}$$

can be manipulated to yield the familiar NPV relationship

$$NPV = (1 - \pi)V_{\mathrm{ND}} + \pi V_{\mathrm{D}} = 0.$$

We start by assuming that the baseline default rate is neither zero nor 1 so that there is at least one default and one non-default. We also assume that fn(k) is not 0 for at least one value of k.

At point k, we are considering the default rate of all firms whose scores fall between percentiles k and $k + \delta$. Let N be the total number of observations, n_k^{D} be the number of defaulters between percentiles k and $k + \delta$ and n_k^{ND} be the number of non-defaulting firms between percentiles k and $k + \delta$. Then:

$$
\begin{aligned}
n_k^{\mathrm{ND}} &= Np(\mathrm{ND}) \cdot \mathrm{tn}(k) \cdot \delta, \quad \text{and}\\
n_k^{\mathrm{D}} &= Np(D) \cdot \int_k^{k+\delta} \mathrm{fn}(\tau)\mathrm{d}\tau
\end{aligned}
$$

since fn(k) is the density of the bads at k and tn(k) is the density of the goods at k (note that tn(k) is a constant).

Therefore, the probability of default, p_k, for firms with a score between k and $k + \delta$ is

$$p_k = \frac{n_k^{\mathrm{D}}}{(n_k^{\mathrm{D}} + n_k^{\mathrm{ND}})}.$$

Substituting, we get:

$$p_k = \frac{Np(D) \cdot \int_k^{k+\delta} \mathrm{fn}(\tau)\mathrm{d}\tau}{N[p(\mathrm{ND}) \cdot \mathrm{tn}(k) \cdot \delta + p(D) \cdot \int_k^{k+\delta} \mathrm{fn}(\tau)\mathrm{d}\tau]}$$

$$(1 - p_k) = \frac{Np(\mathrm{ND}) \cdot \mathrm{tn}(k) \cdot \delta}{N[p(\mathrm{ND}) \cdot \mathrm{tn}(k) \cdot \delta + p(D) \cdot \int_k^{k+\delta} \mathrm{fn}(\tau)\mathrm{d}\tau]}.$$

Now we take the limit of p as δ goes to zero. Applying a special case of the Leibniz Integral Rule, [19] we note that

$$\lim_{\delta \to 0} \frac{\partial n_k^{\mathrm{D}}}{\partial \delta} = N \cdot \lim_{\delta \to 0} p(D) \cdot \mathrm{fn}(k + \delta) = N \cdot p(D) \cdot \mathrm{fn}(k)$$

Now applying L'Hôpital's Rule,

$$\lim_{\delta \to 0} p_k = \frac{p(D) \cdot \mathrm{fn}(k)}{p(\mathrm{ND}) \cdot \mathrm{tn}(k) + p(D) \cdot \mathrm{fn}(k)} \tag{B.2}$$

and

$$\lim_{\delta \to 0} (1 - p_k) = \frac{p(\mathrm{ND}) \cdot \mathrm{tn}(k)}{p(\mathrm{ND}) \cdot \mathrm{tn}(k) + p(D) \cdot \mathrm{fn}(k)}. \tag{B.3}$$

Examining (B.2) and (B.3), we can now rewrite (B.1) as

$$b_k(\mathrm{TN})(1 - p_k)A = c(\mathrm{FN})p_k A$$

where $A = (p(\mathrm{ND}) \cdot \mathrm{tn}(k) + p(D) \cdot \mathrm{fn}(k))$, a normalization term that cancels out.

The profit in default, V_{D} and non-default, V_{ND} are, $b(\mathrm{TN})$ and $c(FN)$, respectively, we get:

$$(1 - p_k)V_{\mathrm{ND}} = p_k V_{\mathrm{D}}$$

or

$$(1 - p_k)V_{\mathrm{ND}} - p_k V_{\mathrm{D}} = 0.$$

Note that uncertainty can be incorporated by transforming p_k into a risk-neutral measure (as discussed in footnote 16).

References

Altman, E.I., 2000. Predicting financial distress of companies: Revising the Z-score and ZETA® Models. New York University, New York.

[19] Douglas Dwyer pointed out the applicability of the Leibniz Integral Rule in this case.

Altman, E.I., Haldeman, R., Narayanan, P., 1977. ZETA analysis, a new model for Bankruptcy classification. Journal of Banking and Finance 1.

Baestaens, D.-E., 1999. Credit risk modeling strategies. Journal of intelligent systems in Accounting, Finance and Management 8, 225–235.

Birdsall, T.G., 1966. The theory of signal detectability: ROC curves and their character.

Birdsall, T.G., 1973. The theory of signal delectability: ROC curves and their character. Cooley Electronics Laboratory, Department of Electrical and Computer Engineering, University of Michigan, Ann Arbor, Michigan. Technical report no. 177.

Duffie, D., Singleton, K.J., 2003. Credit Risk. Princeton University Press, Princeton, NJ.

Granger, C.W.J., Pesaran, M.H., 1999. A Decision Theoretic Approach to Forecast Evaluation, 261–287.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. Peninsula Publishing, Los Altos, CA.

Hanley, J.A., 1989. Receiver operating characteristic (ROC) methodology: The state of the art. Critical Reviews in Diagnostic Imaging 29, 307–335.

Hanley, A., McNeil, B., 1982. The meaning and use of the area under a receiver operating characteristics (ROC) curve. Diagnostic Radiology 143, 29–36.

Hastie, T.J., Tibshirani, R.J., 1990. Generalized Additive Models. Chapman and Hall, New York.

Kealhofer, S., Bohn, J., 2001. Porfolio Management of Default Risk. Moody's KMV, San Francisco.

Pepe, M.S., 2002. Receiver operating characteristic methodology. Raftery, A.E., Tanner, M.A., Wells, M.T. (Eds.), Statistics in the 21st Century. Chapman and Hall/CRC.

Provost, F., Fawcett, T., 1997. Analysis and visualization of classifier performance: Comparison under imprecise class and cost distributions.

Sobehart, J.R., Keenan, S.C., Stein, R.M., 2000. Benchmarking quantitative default risk models: A validation methodology, Moody's Investors Service.

Stein, R.M., 2002. Benchmarking default prediction models: Pitfalls and remedies in model validation. Moody's KMV, New York #020305.

Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. American Association for the Advancement of Science 240, 1285–1292.

Stein, R.M., Jordão, F., 2003. What is a more powerful model worth? Moody's KMV, New York #030422.

Swets, J.A., 1996. Signal Detection Theory and ROC Analysis in Psychology and Diagnostics. Laurence Erlbaum Associates, Mahwah, NJ.