Accurately and Efficiently Measuring Individual Account Credit Risk On Existing Portfolios

By: Michael Banasiak & By: Daniel Tantum, Ph.D.

What Are Statistical Based Behavior Scoring Models And How Are They Developed?

Behavior Scoring Models are statistical based credit scoring models specifically designed to evaluate creditworthiness of existing customers. Behavior Scoring Models are derived from a statistical analysis of individual account credit performance. The purpose of the statistical analysis is to find the most predictive set of data elements that separate the good credit risks from the bad credit risks. Then the data elements are weighted statistically typically using a multivariate maximum likelihood estimation technique and an assumed distribution of risk, such as the Logistic distribution, which is particularly well suited for the good vs. bad credit risk class of problems.[1]

In the credit industry, Behavior Scoring Models, and credit scoring models in general, are confused with “Expert Systems” or “Heuristic Rules Based Systems” that use the credit manager’s own experience to choose the data elements and create the data weights or rules to automate the decision process. Basically, Expert Systems replicate in computer program code the credit manager’s data steps in evaluating a particular account that would be as close as possible to a manual credit review. Therefore, Expert Systems provide speed efficiency to the credit evaluation process, by minimizing credit analyst intervention on routine transactions. However, an Expert System’s weaknesses are typically its inability to bring greater accuracy, and effectiveness to the credit evaluation process.

The Behavior Scoring Model being an automated tool provides the analyst instant information, and being a more accurate credit risk tool increases the effectiveness of the credit analyst’s review time. Compared to an Expert System, accuracy and effectiveness, or what can be called decision efficiency, are the strengths of credit scoring and Behavior Scoring Models. The greater accuracy of credit scoring and Behavior Scoring Models comes from the power of mathematics being able to analyze hundreds of credit risk data elements to find the most predictive set of data elements and then optimally weighting the data elements to maximize the model’s predictiveness. The Behavior Model is then able to most accurately identify low risk accounts, without a need for analyst review, freeing up that time for the analyst to spend on the higher risk accounts with more difficult decisions. The predictive power of statistical models is best understood by reviewing the development process and by understanding its contrasts with Expert Systems.

A Major Differences Between Behavior Scoring Models and Expert Systems

Behavior Scoring Models are developed by statistically analyzing historical credit performance on a cohort of individual accounts. Like Expert Systems, Behavior Scoring Models must

[1] The appendix details the statistical foundation of Behavior Scoring Models assuming the Logistic distribution.

replicate the decision process of the credit manager. But, the most notable differences in developing a Behavior Scoring Model are:

1) A Historical Statistical Analysis of Account Payment Behavior - The Behavior Scoring Model is based upon a historical statistical analysis of actual credit decision data and credit performance outcomes. Credit performance is typically measured 6 to 24 months from the credit decision date and is classified into good credit versus bad credit performance. For example, bad credit performance might be defined as a 90+ day delinquency, loss, or bankruptcy, and good credit performance would be defined as not meeting the 90+ day delinquency criteria. Longer performance windows, such as 12 months, are typically required for minimizing loss risk of credit transactions and shorter performance windows meet the needs of internal collectors prior to 3rd party collections.

2) Statistics Determines The Data Elements That Are Predictive - The historical multivariate statistical analysis of the credit decision data, which could entail accounts receivable, financial statement, or credit bureau data, determines which data elements are most predictive. This process usually finds many of the traditional credit risk data elements, that would also be found in an Expert System, to be predictive of credit risk with the expected economic relationships. However, the power of statistics brings added value by uncovering predictive data elements that may be less obvious to the credit manager and by eliminating traditional data elements that are found to be less predictive than others.

3) Statistics Assigns Optimal Weights To Maximize Predictiveness - In developing a statistical based Behavior Scoring Model, the credit manager does not choose the model weights. Statistical maximum likelihood estimation method assigns the weights mathematically to optimize the predictiveness of the model. In other words, this method maximizes the separation of the good credit risks from the bad credit risks, which allows for the accurate identification of credit risk. Based upon the mathematical properties of the statistical maximum likelihood technique, the Behavior Scoring Model will always outperform the Expert System or in a rare special case at least match its performance. Given this greater accuracy and minimum criteria, creditor firms are better off applying the Behavior Scoring Model technique.

4) Predictive Power Is Proven Through Validation - The predictive power of the Behavior Scoring Model is proven through model validation. The validation is conducted by applying the Behavior Scoring Model formula (i.e., data elements and weights) to the sample the model was developed upon, called the Development Sample, and to another sample of individual accounts not analyzed in the model development process, called the Holdout Sample. Then model accuracy is tested by ranking (i.e., sorting) from the highest probability of achieving the bad credit performance definition to the lowest probability. The results on both the development and holdout sample can be compared based upon ten percentile groupings and by bad credit performance capture rates (i.e., cumulative distribution of bad credit performance accounts). If the model results are statistically similar on the Holdout Sample as compared to the Development Sample, this proves that the model works effectively on another sample. Validation examples are provided in Figures 4.3 and 4.4 of the Behavior Model for Collections, Case Study – Validation Methods.

5) Credit Risk Performance Is Measured By The Behavior Scoring Model – Based upon the

Development and Holdout Samples, credit risk performance tables are created for credit management to understand the expected credit performance of the score. These Behavior Scoring Model performance tables are used to make score cut-off point decisions to automate decisions by various credit risk levels, which provides great flexibility to the credit manager to change cut-off points very quickly if necessary. Since the output of the Behavior Scoring Model is the probability of bad credit performance and the probabilities are very non-linear, generally the probability is transformed into a 1 to 100 score that maintains the ranked order. Then a table of scores can be created that shows expected credit performance based upon actual accounts, such as marginal 90 plus day delinquency rates, cumulative bad credit performance capture rates, bad credit performance rates for any proportion of accounts above score, at score, or below score. An example of a Performance Management Table is provided in Table 3.1.

The Output of Behavior Models – Scores, Performance Management Tables and How They Are Used

The output of a behavior model is the probability that an on-going account will become severely delinquent, or experience a loss or bankruptcy (behavior probability). The behavior probability can be used for many important credit decisions:

- Identify accounts that are very likely to experience severe delinquency, loss or bankruptcy over the near future, even those with no past history of delinquency, and allow pre-emptive action to avoid the severe delinquency or loss

- Identify accounts that are very likely not to experience severe delinquency, loss or bankruptcy over the near future, even those with a frequent past history of delinquency, and therefore free up working resources to focus on those accounts with a greater likelihood of severe delinquency

- Identify accounts for which either authorization of new business or a credit line increase is very likely to increase credit risk beyond a tolerable level, and therefore subject the request to further review and/or automatic declination.

The power of the behavior probability is maximized when it is used in conjunction with a Performance Management Table (PMT). The Performance Management Table provides a complete picture of the distribution of expected risk for a cohort of accounts. Based on ones tolerance for risk, and other business objectives, the PMT helps credit managers decide where to "draw the line" for acceptable risk. Where that line is drawn determines what accounts are to be worked by an analyst, as is the case in identifying accounts for collections, or what account requests are subject to further review, or automatic declination, as is the case for a request of additional business or increased credit line.

Before we discuss the specifics in the use of a PMT let's refer to a hypothetical, but realistic, PMT as shown in Table 3.1.

Table 3.1
Performance Management Table
Behavior Scoring Model

Risk Distribution by Score

Score	Cumulative % of Total	Cumulative % of Total Expected Bads	Expected Marginal Bad Rate	Expected Bad Rate Score and Above	Expected Bad Rate Score and Below
1	0.70%	11.76%	55.20%	4.94%	82.80%
2	1.13%	17.14%	41.56%	4.38%	75.06%
3	1.36%	19.70%	36.07%	4.13%	71.46%
4	1.51%	21.25%	32.83%	4.01%	69.18%
5	1.64%	22.39%	30.44%	3.95%	67.41%
6	1.75%	23.39%	28.69%	3.89%	65.82%
7	1.86%	24.29%	27.14%	3.84%	64.35%
8	1.97%	25.16%	25.89%	3.80%	62.93%
9	2.07%	25.91%	24.70%	3.77%	61.68%
10	2.19%	26.75%	23.64%	3.74%	60.29%
11	2.29%	27.49%	22.76%	3.69%	59.06%
12	2.40%	28.22%	21.90%	3.66%	57.87%
.	.	.	.	.	.
.	.	.	.	.	.
44	7.84%	50.15%	8.89%	2.69%	31.55%
45	8.11%	50.86%	8.66%	2.67%	30.92%
46	8.38%	51.56%	8.43%	2.64%	30.33%
47	8.68%	52.32%	8.19%	2.61%	29.70%
48	8.91%	52.87%	7.97%	2.58%	29.25%
49	9.51%	53.52%	7.76%	2.55%	28.71%
50	10.00%	55.00%	7.15%	2.48%	27.75%
51	10.21%	55.15%	7.02%	2.49%	27.56%
52	10.32%	55.57%	6.95%	2.46%	27.11%
.	.	.	.	.	.
.	.	.	.	.	.
89	46.79%	90.13%	1.31%	0.95%	9.50%
90	50.01%	90.96%	1.18%	0.92%	9.14%
91	53.19%	92.25%	1.04%	0.87%	8.55%
92	57.84%	93.58%	0.94%	0.81%	7.98%
93	62.39%	94.69%	0.80%	0.75%	7.49%
94	69.82%	96.21%	0.67%	0.69%	6.80%
95	75.32%	97.19%	0.59%	0.62%	6.36%
96	87.03%	98.84%	0.46%	0.56%	5.60%
97	95.07%	99.67%	0.34%	0.44%	5.16%
98	99.26%	99.97%	0.24%	0.33%	4.97%
99-100	100.00%	100.00%	0.13%	0.20%	4.94%

The columns of the PMT provide the following information:

Score – Provides a numerical measure of risk. It is a direct transformation of the behavior probability, and has been calculated here to range from 1, maximum risk, to 100, minimum risk, with risk decreasing as the score increases.

Cumulative % of Total – The percent of all accounts that are covered up to and including any score level. A score of 50 covers the riskiest 10 percent of accounts in the portfolio.

Cumulative % of Total Expected Bads – The percent of all accounts that are expected to experience bad credit performance up to and including that score level. A score of 50 captures 55% of all expected bads in the portfolio

Expected Marginal Bad Rate – The percent of accounts that are expected to experience bad credit performance within any score level. For accounts that score exactly 50, not 51 and not 49, 7.15% of them are expected to experience bad credit performance.

Expected Bad Rate Score and Above – The percent of accounts at, or above, score that are expected to experience bad credit performance. The expected bad rate for accounts that score 50 or above is 2.48%.

Expected Bad Rate Score and Below – The percent of accounts at, or below, score that are expected to experience bad credit performance. The expected bad rate for accounts that score 50 or below is 27.75%.

The optimal use of the PMT can differ by the credit question; identifying an account for internal collection work, or screening out accounts that have requested credit authorization for additional business or an increase in credit line for further review. We will now take a look at the use of a PMT use for each credit question.

Use of Performance Management Tables For Repeat Credit Authorizations or Credit Line Increases

The first course of action in the use of the PMT is to decide two different score cut-off levels. The two score cut-off levels result in three courses of action, and they are:

- Approval cut-off point, at score and above the request is approved
- Decline cut-off point, at score and below the request is declined
- In-between the cut-off points the request is subject to further review

The score choice is unique to a company's risk tolerance and business objectives. The company credit manager should not make such a choice in isolation. As is readily seen in the PMT, the higher the score chosen for decline the more business foregone. In the extreme one could use their PMT to drive the delinquency rate to close to zero, but a consequence is that revenue is likely to be driven to unacceptable levels. The Score choice should be based, not only on portfolio credit quality objectives, but also with the consideration of business revenue and profit objectives. The Score choice will then reflect what is in the company's overall best interest. A company that is aggressively trying to grow business will set a lower score for approval than a company that is focused on reducing their rate of severe delinquency.

To make clear the interpretation and value of the PMT, let's work through a hypothetical example. After the proper decision review process, a score of 90 and above was chosen for automatic approval, a score of 50 and below was chosen for automatic decline, and any score in-between was subject to further review. According to the PMT of Table 3.1 the company can expect to automatically approve 50% of the requests with an expected bad rate less than 1%, and automatically decline 10% of the requests to avoid an expected bad rate of 28%. The requests that score between 50 and 90 will make up the remaining 40% the requests, and are subject to further review.

Use of Performance Management Tables For Collections

The first course of action in the use of the PMT to support collections decisions is to decide a score cut-off level that will generate two different courses of action. The score cut-off level and two courses of action are:

- Score and above no collection activity is required
- Below score the account is pulled for collection activity consideration

The score choice should be similar to the process already described in choosing scores for authorization; input and buy-in across organizations with the score chosen consistent with overall business objectives.

Referring to Table 3.1 we see that with a chosen score of 45, the company can expect to pull about 8% of its accounts for activity. However, it will be focusing its resources and efforts on over 50% of the accounts, that without remediation would be expected to experience severe delinquency in the near term.[2] If a company's desire was to more aggressively reduce severe delinquency it may set the score at 50, but note increased resources may be needed to handle the increased work load (about 10% of the accounts).

The use of the PMT for collection activity prioritizes accounts to be worked based on the probability of near term severe delinquency loss or bankruptcy. A more powerful prioritization would take into account the magnitude of risk exposure, given the probability. As it stands, an account with an outstanding balance of $100 and a score of 35 may receive as much resource attention as an account with the same score but an outstanding balance of $10,000. The authors therefore recommend a two-stage prioritization approach to maximize the effectiveness of a company's collections resources. The first stage is to choose the cut-off score as already described. The second stage is to subject all the accounts under the score cutoff to a second calculation. The calculation is to multiply the known behavior probability by the known

[2] Note that while the concepts will not differ, the actual values of the Performance Management Tables will differ depending on the Behavior Model. The reference to a single table is for simplicity.

outstanding balance at time of score. The calculation provides a measure of expected dollar risk exposure. After the calculation, re-rank by the expected dollar risk exposure. The prioritization of resources for each account can now take into account, not only the behavior probability of bad credit incidence, but also the dollars that are at risk.

Case Study - Models For Credit Authorization or Due Diligence Review and Performance Validation

Most companies are faced with making day to day credit authorization decisions for repeat business of existing customers. Many companies also do year end portfolio reviews, also known as annual due diligence reviews. The authors have built Behavior Models for exactly those needs. In the case study that follows we will discuss three Behavior Scoring Models that were built for a financial services company. The distinction among the models is the use of bureau data, and if bureau data is used, the source of bureau data. The bureau data comes from the two major commercial credit data providers. We will refer to the commercial data providers as Bureau 1 and Bureau 2. The models using bureau data also use accounts receivable data. The accounts receivable model (AR model) as its name implies, uses solely accounts receivable data.

For the following reasons three models were built:

1) Bureau data elements provide additional information - more data elements to optimize predictiveness is better than fewer
2) Two different bureaus were used to maximize the hit rate
3) The use of two different bureaus does not guarantee complete coverage

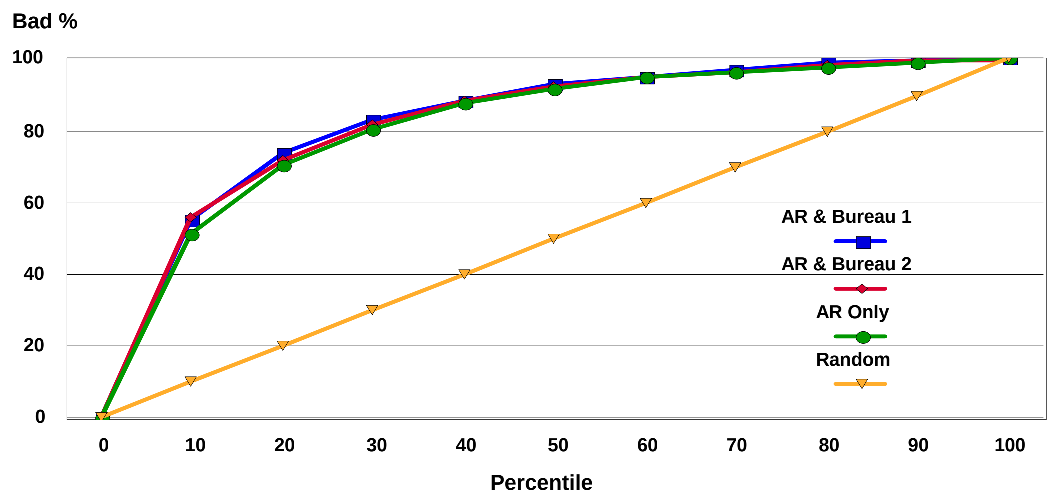
The models were built to produce monthly scores. The use of monthly scores means credit authorization decisions will be based on very “fresh” data, as well as provide a tool that effectively supports a more timely due diligence process, e.g. quarterly due diligence.

The performance and effectiveness of the models can be seen and compared with analysis of each model’s Credit Screening Effectiveness and Separation Effectiveness. To measure credit screening effectiveness we use the Behavior Model to calculate the probability of bad credit performance. We then sort the accounts by descending probability, resulting in a ranking of accounts per the probability, from greatest risk to least risk. After the sort we can say the top 10% of the sorted accounts are in the top 10th percentile of risk, the top 20% of the sorted account are in the top 20th percentile of risk, etc. The Behavior Model’s Credit Screening Effectiveness is then identified by the proportion of known bad captured at any percentage point of the ranked distribution of risk.

Figure 4.1 shows the Credit Screening Effectiveness of each model. Figure 4.1 provides percentiles along the x-axis and the percent of known bad captured along the y-axis (Bad %). To make clear how to read Figure 4.1, consider the points on the AR & Bureau 1 curve, AR & Bureau 2 curve, and the AR Only curve. At the 20th percentile the AR & Bureau 1 model captured 74% of the future bads, the AR & Bureau 2 model captured 72% of the future bads, and the AR Only model captured 71% of the future bads.

Staying with the 20th percentile point, the Credit Screening Effectiveness literally says, that if the customer had used these models at time of credit evaluation they would have predicted over 70% of the bads before the severe delinquency even occurred. This result is very powerful. Of course, such knowledge is used to change the future so that many of the same 70% would not have had the opportunity to go bad, or that if the bad had occurred, actions would have been taken to minimize the impact. The result clearly illustrates the value of a Behavior Model with strong screening effectiveness.

Figure 4.1
Measuring Credit Screening Effectiveness
Behavior Model For Repeat Credit Authorizations



Case Study - Performance and Data Source Comparison

Another important result shown in Figure 4.1 is the power of AR data. The AR Only model shows the predictive power of accounts receivable data. Based on the authors' experience in building Behavior Models, the screening effectiveness power of the AR Only model shown here can be considered representative. The AR Only model screening effectiveness is not as great as the models supplemented with bureau data, but the difference is not "large". However, keep in mind two caveats. First, this is a case study, and you will find smaller or larger differences on a portfolio by portfolio basis. Second, what constitutes a large difference is at the discretion of each company, as determined by the company's business objectives.

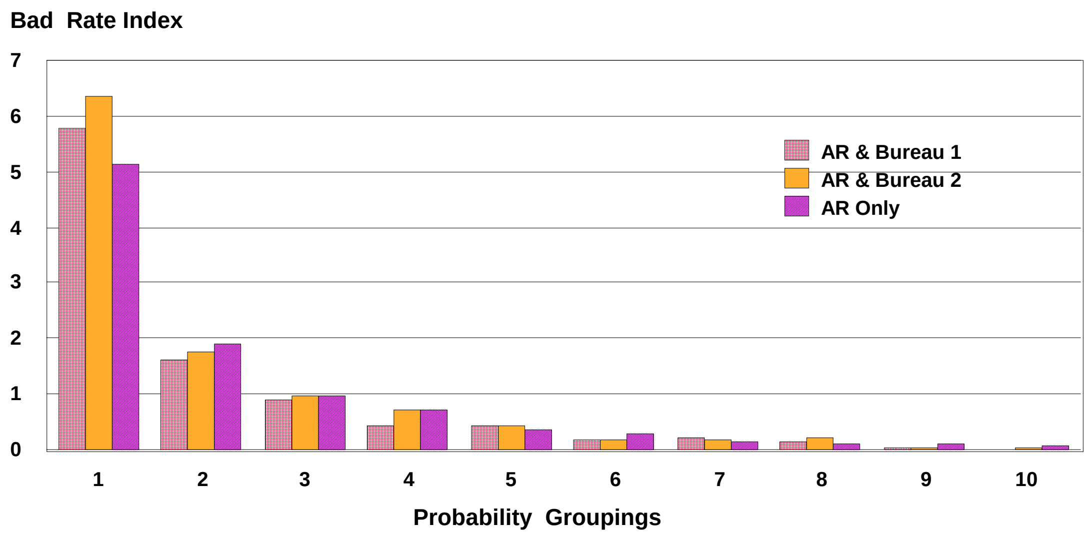
The method to identify the Credit Separation Effectiveness of a model starts with the same process to identify screening effectiveness: a ranking of accounts with known good or bad credit behavior by descending probability of bad. The ranked accounts are then sorted into mutually exclusive probability groups. The group with highest probability range should have the highest bad rate (the known bad accounts in the group divided by the total number of accounts in the group), the group with the second highest probability range should have the second highest bad rate, and so on.

Figure 4.2 shows the Credit Separation Effectiveness of each model. A simple index is substituted for the bad rate within each of the 10 groups to more accurately reflect the absolute magnitude of the separation effect for each model. The index is the ratio of the bad rate within each group divided by the bad rate for the total sample. An index value of one would indicate that the bad rate for that group was equal to the bad rate for the overall sample. The results shown are consistent with models that have demonstrated strong separation power. Accounts in Group 1, the highest risk group, have a bad rate between five and six times greater than the bad rate for the entire sample. For the AR & Bureau 2 model, the index is 6.3; for the AR & Bureau 1 model it is 5.8; and, for the AR Only model the index is 5.1. At the other extreme of the probability distribution are the accounts in Group 10, the lowest risk group. Here, the maximum value of the bad rate index is 0.10, reflecting a bad rate that is 10% that of the overall sample. Another desirable finding is the almost monotonic decrease in the Bad Rate Index moving from Group 1 to Group 10. These results demonstrate that rank ordering accounts by the probability of going bad, as estimated by each respective model, yields a strong separation effect, in separating good accounts from bad accounts.

Figure 4.2

Measuring Credit Separation Effectiveness

Behavior Model For Repeat Credit Authorizations



Case Study - Performance and Data Source Comparison

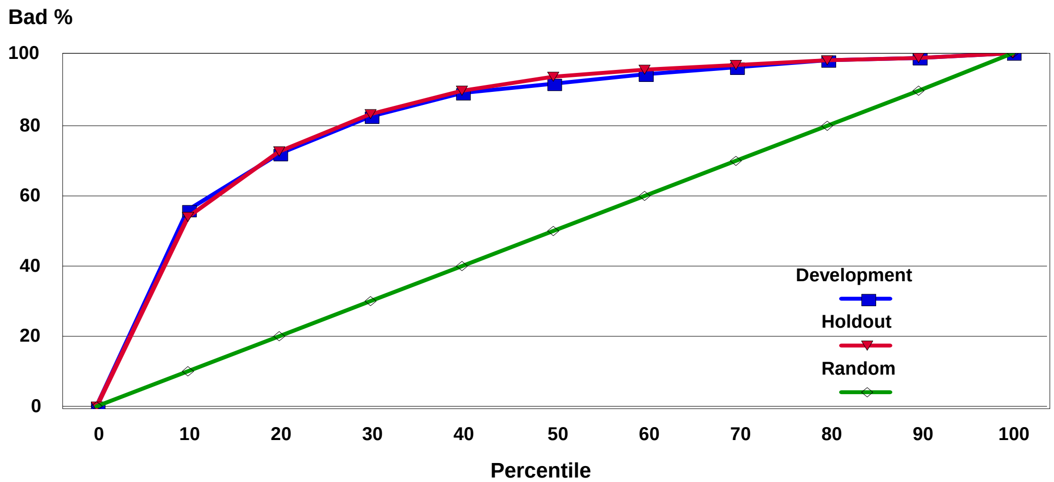
Case Study – A Model For Collection Action Decisions and Model Validation

In the previous section we focused on the overall performance and comparison of the overall performance of three Behavior Models. Here our focus is on Model Validation. By validating a model we mean comparing the model's predictive ability to screen and separate bads based on the same sample of accounts used to the develop the model (Development Sample), with the model's predictive ability on a sample not used in the model development process (Holdout Sample). Model validation is a very important step in Behavior Modeling. If the screening and separation results are quite similar, you have evidence that the model will work outside the "laboratory" and in the real world. If the Holdout Sample results are much weaker, you have a warning that the model may not be valid.

To illustrate the use of validation testing Behavior Models, we present the case of a Behavior model built for a company providing net 30 day financing to the general mix of the business universe. The Behavior Model was based on the customer's historical account receivable data only, and was developed especially for collection action decisions. A Behavior Model for Collections is built to predict over a relatively short performance period. Here the model is predicting a 91 plus day delinquency, loss or bankruptcy within 6 months. Behavior Models for collections are used to identify accounts that are very likely to experience immediate to near term difficulty in meeting their current outstanding obligations. On the other hand, the Behavior Models for credit authorizations are predicting the performance of additional activity that will take time to reflect its contribution to the overall credit worthiness of the account. The validation tests we describe here are applicable for all Behavior Models, and should be performed for all Behavior Models.

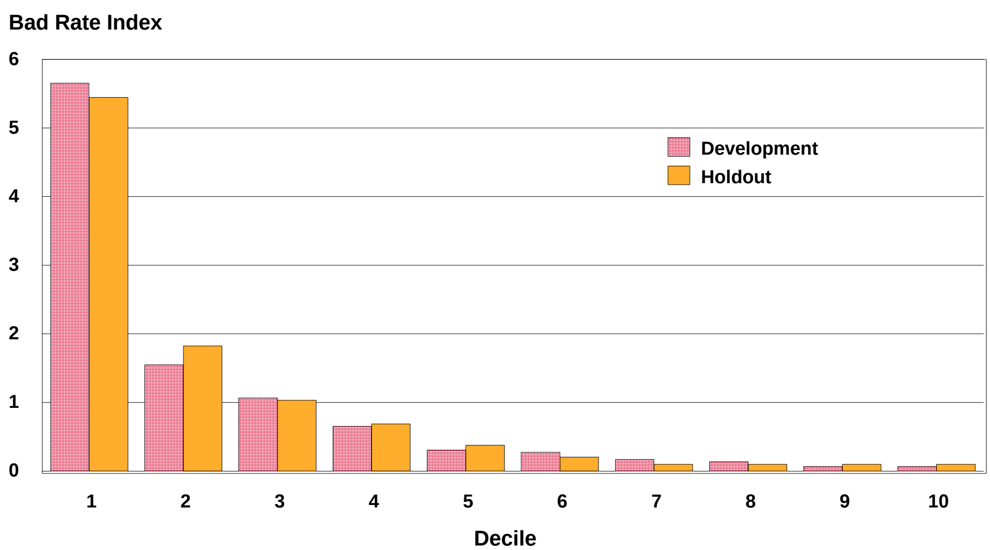
Figure 4.3 and Figure 4.4 compare the Credit Screening Effectiveness and Credit Separation Effectiveness of the Development Sample with the Holdout Sample. The results shown illustrate the Behavior Model's ability to successfully screen and separate outside of its own Development Sample. The screening effectiveness lines practically appear to be one line throughout the whole risk distribution. At the 20th percentile 72% of the bads are captured in the Development Sample and the Holdout Sample. As was the case for credit authorization models, the Credit Screening Effectiveness literally says, that if the customer had used these models at time of collection evaluation they would have identified over 70% of the bads before the severe delinquency, loss or bankruptcy even occurred. In the case of collection decisions, such knowledge would have been used to separate these accounts for collection activity to prevent the bad behavior, or to minimize its impact. The separation effectiveness shown in the Development and Holdout samples is also very close, with the bad rate in the highest risk group across both samples, about 5.5 times greater than the bad rate for the total sample. The marked closeness of the bad rate by risk group is evident throughout the entire risk distribution. In summary, the validation method case study proves that this Behavior Model works very well on another sample.

Figure 4.3
Measuring Credit Screening Effectiveness
Behavior Model For Collections



Case Study - Validation Methods

Figure 4.4
Measuring Credit Separation Effectiveness
Behavior Model For Collections



Case Study - Validation Methods

Conclusion:

Behavior Scoring Models use the power of statistics to measure creditworthiness on existing accounts. The results are proven to be real world effective through statistical validation on a sample untouched in the development of the model. The use of Maximum Likelihood Estimation gives the Behavior Scoring Models an unfair advantage in outperforming Expert Systems in predicting credit risk. Credit risk performance management tables allow the credit manager to optimize the use of the Behavior Scoring Model, by providing a summary of risk measurement for the portfolio.

The case studies show that accounts receivable data is the most important input to a Behavior Scoring Model in predicting the credit risk from both a repeat credit authorizations, and from a collections perspective. Also, the credit authorization case study shows that commercial credit bureau data incrementally increases credit risk prediction accuracy, when added to the accounts receivable data in a Behavior Scoring Model.

The outcome of using Behavior Scoring Models is greater accuracy, effectiveness, and speed and decision efficiency for the credit decision process.

Appendix

Models of Qualitative Choice and Behavior Models

Models of qualitative choice provide a valuable understanding of when and why a particular choice is made. Some examples of choice are; to take the car, bus, or subway to work; to respond or not respond to a mail solicitation; to be severely delinquent or not be severely delinquent in paying a bill. Models of qualitative choice reveal the drivers of choice and provide the probability that a particular choice will be made.

Our purpose is to build qualitative choice models that provide a very accurate measure of predictive credit risk for an on-going account. A measure of predictive credit risk is the probability that an account will choose to be a credit risk, such as the account choosing to be severely delinquent, or go to loss or bankruptcy. The high accuracy of the probability is accomplished by exploiting a rich set of information specific to the account, and highly correlated to the future performance of that account. Predictive credit performance information for on-going accounts includes, but is not limited to:

Historical accounts receivable credit performance
Characteristics of the lender/borrower relationship
Current values of various financial measures
General trade information from commercial and consumer bureaus

The combination of the question at hand i.e., what is the probability that an on-going account will be a credit risk or not a credit risk, and the use of information about the account's own past payment performance uniquely defines the qualitative choice model as a behavior model.

How is a probability of credit risk obtained using predictive credit performance information? Behavior models are simply mathematical functions that take as inputs the predictive credit performance information and provide as output the probability of credit risk. A most common form of the mathematical function used is the logistic regression equation. The logistic equation is:

$$F(s) = \exp(s) / (1 + \exp(s)). \qquad (1)$$

Where exp = exponential mathematical operator.

F(s) is read as a function of the value of the score or parameter, *s*. A property of equation (1) is that, as *s* ranges from negative infinity to positive infinity, F(s) will increase from zero to unity. And so, F(s) meets a necessary condition for a probability function in that it is bounded between zero and one. Another property of F(s) is that its graph is S-shaped. A property of the independent parameter *s*, is that the probability of its value is symmetrically distributed about zero, meaning there is an equal probability (50%) that its value will be greater than or less than zero.

When we interpret F(s) as the probability that an account will be severely delinquent we have:

$$\text{Probability} = \exp(s) / (1 + \exp(s)) \qquad (2)$$

Where Probability is read as the probability an account will be severe delinquent.[3] As already indicated this probability is a function of the value of s. We also know that as s increases, the probability of severe delinquency increases. What determines the value of the parameter s? The logistic equation for probability identifies parameter s as an easily understood and straightforward measure of risk. As it turns out the parameter s in the logistic equation is the natural logarithm of the odds (log odds).[4] For the problem at hand the ratio of Probability to one minus Probability tells us the odds that the account will be a credit risk versus will not be a credit risk. Assuming Probability follows a logistic equation we obtain:

$$\text{Odds} = \text{Probability} / (1 - \text{Probability}) = \exp(s) \qquad (3)$$

When you take the natural logarithm of the odds you obtain:

$$\text{LOdds} = \text{Ln}(\text{Odds}) = \text{Ln}(\exp(s)) = s \qquad (4)$$

[3] Note that since an account will either be severely delinquent or not severely delinquent the equation to determine the probability that an account will not be severely delinquent is equal to 1 – Probability, or 1 / (1 + exp(s)).

[4] The logistic equation is chosen to model probability because of the parameter s straightforward interpretation as the log odds, and the easily understood relationship of odds to probability for the two choice case. For an intuitive justification of qualitative choice models that allow any continuous probability distribution function defined over the range of negative infinity to positive infinity see Aldrich and Nelson, 1984, pp. 35-37.

Therefore the probability to be severely delinquent can be modeled by simply postulating a model for the odds to be severely delinquent versus not severely delinquent.

We can model our LOdds as:

$$LOdds = B_0 + B_1{*}X_1 + B_2{*}X_2 + B_3{*}X_3 + B{*}X_4 + B_5{*}X_5 \ldots\ldots.+ B_n{*}X_n \qquad (5)$$

Where:
B_0 = Intercept term
X_i = Predictive Credit Data Element
B_i = Parameter for data element X_i

Note that each variable in equation (5) has a parameter. Utilizing historical values of your accounts' credit performance and predictive credit performance variables, and the proven statistical estimation method, maximum likelihood estimation, generates highly accurate parameter estimates. The estimated parameter for each variable provides a "weight" of that variable's contribution to an account's odds of becoming severely delinquent. Replacing the "weights" and variables of equation (5) with the estimated "weights" and current values for the predictive credit performance variables of any account, we are able to obtain the current probability of any account to become severely delinquent:

$$Probability = \exp(LOdds) / (1 + \exp(LOdds)) \qquad (6)$$

References

Aldrich, J.H. and F.D. Nelson (1984) Linear Probability, Logit, and Probit Models. Sage University Papers: Quantitative Applications in the Social Sciences, 07-45. Beverly Hills, CA: Sage.

Michael Banasiak is President and Daniel Tantum, Ph.D. is Vice President of Analytical Services at Predictive Business Decision Systems. The authors can be reached via e-mail at mbanasiak@pbdsinc.com or dtantum@pbdsinc.com.

Credit Research Foundation
8840 Columbia 100 Parkway
Columbia MD 21045
410-740-5499