

## Inferring Behavior on Rejected Credit Applicants-Three Approaches

Gregg Weldon, Sigma Analytics & Consulting, Inc., Atlanta, GA

### ABSTRACT

Ever since statistical modeling gained widespread acceptance in the credit industry, portfolio managers have struggled with how to build predictive models on their true, through-the-door population. In many cases, one third or more of the applicants that cross a lender's threshold are denied credit. Since some of any lender's approvals fail to perform as agreed, it's reasonable to assume that some of the lender's rejected applicants would have likewise paid satisfactorily. Any through-the-door risk model, then, must take all applicants into account, not just the ones whose performance is a known quantity. This paper focuses on three commonly-used methods of inferring behavior on these rejected applicants - Cohort Performance, Parceling, and Augmentation. Each method has its own strengths and weaknesses.

### MODELING ON KNOWN "GOODS" AND "BADS"

In any credit modeling procedure, it's tempting to stay as close as possible to "known" information. Because of the rigid regulatory yet highly competitive environment most creditors operate in, any theoretical or esoteric solutions generally take a backseat to a more quantifiable, dollars-and-cents answer. In some cases in the past, that has translated into designing scorecards to predict the overall population that are built only on a lender's known Goods (applicants who were approved and paid as agreed) and Bads (applicants who were approved and paid slowly, if at all).

The problem with this approach can be illustrated as follows. Assume that a small bank only approves applicants who have never had a car repossessed. The bank feels that repossession is a signal of poor credit performance. Indeed, the credit industry is founded on the belief that past behavior is predictive of future performance. One day, the son of the bank's largest depositor arrives requesting a car loan, despite the fact that his last car was repossessed. He states that he has learned from his mistakes and will never go delinquent again. Because of his father, the bank overrides its policy and makes the loan. The son is as good as his word

and he pays the loan off as agreed. If a scorecard was built on this bank's known Goods and Bads, it's conceivable that positive points would be assigned for having a past repossession. The scorecard only experienced one repossession in its data and it was a Good. When this model is used on the bank's through-the-door population, loan losses could be expected to increase very quickly.

The above example is extreme, but it points out a problem that does occur with alarming frequency for creditors. The only way to build an accurate scorecard for use on the public at large (or at least the creditor's portion of the public at large) is to rely on approved accounts AND rejected applicants. Leaving the "known" credit world, then, is essential.

### COHORT PERFORMANCE

In the best scenario, a creditor may turn down an applicant for a car loan, watch that applicant apply for and receive the same loan with a creditor across the street, and find out later how he paid the competition. In some cases, this method of Cohort Performance is available to creditors. The major credit bureaus (Equifax, Trans Union, and Experian) have access to the credit files of the great majority of the nation. If one bankcard issuer refused credit to an individual, it's possible to see how that individual performed with other bankcard issuers over a certain period of time. If performance with bankcards is too narrow a definition, performance with all trades is also possible.

Obviously, Cohort Performance is a very powerful method of obtaining performance information on rejected applicants. There are, however, some drawbacks. These include the cost of purchasing this information, the time it will take to receive it, and the overall quality of the data itself. The biggest drawback, however, is in the assumption that the applicant you turned down was able to obtain credit elsewhere. Many applicants have credit histories so bad that getting approved is either nearly impossible or extremely expensive. Lenders charge applicants based on risk, so the biggest credit risks will pay the highest rates, have the most collateral requirements, and be the first borrowers the collection agencies go

after in the event of default. Any credit behavior they have may be skewed because of these stringent requirements. Many times, the rejected applicants with the worst credit histories (the ones that most lenders need to infer behavior on the most) are the least likely to have Cohort Performance on which to infer.

## PARCELING

Parceling is a method of reject inferencing that avoids the problems associated with Cohort Performance. Parceling segments a creditor's Goods, Bads, and Rejects by some generic or custom risk score and then "infers" behavior on the Rejects at the same proportion of Goods and Bads as the approved applicants.

In the SAS® system, a frequency is run using a Score by some performance indicator, such as BGR, where Bad=0, Good=1, and Reject=2. In the example below, PROC TABULATE was used:

SCORE	BAD #	GOOD #	BAD %	GOOD %	REJECTS
LOW-199	16	0	100%	0%	214
200-299	13	68	16.05%	83.95%	262
300-399	68	564	10.76%	89.24%	665
400-499	44	1512	2.83%	97.17%	933
500-HIGH	28	1964	1.41%	98.59%	285

At this point, we could create variables called UNI, BAD, GOOD, REJECT, and ACCEPT for use in Parceling. UNI is a random number generator created from:

```
UNI = (ranuni(seed));
BAD = (BGR = 0);
GOOD = (BGR = 1);
REJECT = (BGR = 2);
ACCEPT = 1 - REJECT1;
```

The Rejects are then parceled among the Bads and Goods at the same rate as the approvals. For example, all 214 Rejects that scored from LOW to 199 would be made Bads. 16.05% of the 262 Rejects that scored from 200 to 299 would be made Bads, with the other 83.95% becoming Goods.

```
IF (BAD = 1) OR
((SCORE <= 199) & (REJECT = 1) & (UNI <= 1.0))
OR
((200 <= SCORE <= 299) & (REJECT = 1) & (UNI <=
.1605)) OR
```

*more lines of code*

```
THEN DO;
    BAD1=1;
    GOOD1=0;
    REJECT1=0;
    ACCEPT1=1;
    END;
ELSE DO;
    BAD1=0;
    GOOD1=1;
    REJECT1=0;
    ACCEPT1=1;
    END;
```

At this point, modeling is done on BAD1 and GOOD1 rather than on BAD and GOOD because the new variables now also contain Rejects.

Another alternative in Parceling is to parcel a higher proportion of Rejects to the Bads at each interval than was calculated above. This is to reflect the belief that these Rejects were denied credit for a reason and are not really as much "like" the similarly scoring Bads and Goods as the score used for this analysis seems to reflect.

Although Parceling is better able to infer behavior for the worst credit applicants, something that Cohort Performance was unable to do, it does have its own weaknesses. First, parceling relies on the utilization of a good Score on which to parcel. A score that is unable to separate Goods from Bads adequately, or was built on specious data, can have an adverse effect on the parceling of Rejects. Parceling is considered a conservative method of reject inferencing because portfolios with low Bad rates would result in having few Rejects becoming Bads and many Rejects becoming Goods. This may actually "water down" performance more than some other methods. Also, because the parceling is performed on a single score alone, it is limited in its ability to identify and correct for truly substandard applicants as well as Augmentation, the next method of reject inferencing.

## AUGMENTATION

Augmentation is a more complex but more complete method of reject inferencing. Rather than relying on a single score to determine how Rejects are counted, Augmentation takes a multi-dimensional approach. Augmentation can be divided into two parts: reclassification and reweighting. As mentioned earlier, one drawback to Cohort Performance is the possibility of overlooking the absolute worst credit that any model will need in order to identify all aspects of credit history.

Augmentation begins with this step, called reclassification.

There are some credit attributes that are so bad (i.e. representative of future delinquency) that most creditors will reject applicants outright for them. These include prior bankruptcies, charge-offs, trades that are currently 90 days past due or worse, etc.. The first step in Augmentation is to look at the creditor's data and see which credit attributes, or variables, the creditor considered deal-breakers.

Let's assume that a lender has a portfolio with a 5% delinquency rate. One of the attributes that this creditor rejects highly on is DERPR (number of derogatory public records: liens, charge-offs, and garnishments). In fact, 90% of all applicants with DERPR >= 1 who applied for credit with this lender were turned down. This may be a good candidate for use in reclassification. Although 90% were turned down, the creditor approved 10%. These applicants must have had some overriding aspect to their application that the creditor felt that they would be the exception to the rule, the cream of the crop. Examination of the Bad rate for approved applicants who had DERPR >= 1 will give an indication of how well these overrides performed. In this example, those approved applicants have a bad rate of 11.2%. This is much higher than the overall population bad rate of 5%. This indicates that even the best people with this attribute are much more likely to go delinquent. It may be reasonable to assume that had the other 90% of those applicants with DERPR >= 1 been approved, they would have gone delinquent.

```
IF (BAD = 1) OR
  (( REJECT = 1) & (DERPR >= 1))
THEN DO;
  BAD1 = 1;
  GOOD1 = 0;
  REJECT1 = 0;
  ACCEPT1 = 1;
END;
```

With reclassification completed, it is once again necessary to use a Score to segment the data. This time, however, the score will be used on Rejects versus Accepts rather than on Goods and Bads. Any score (custom or generic) that is able to score Rejects low and Accepts high will do, the greater the separation, the better. One method of testing this separation is the Kolmogorov-Smirnov (KS) test, an industry standard.

Unlike Parceling, no Rejects will be brought into the sample and made a Good. In fact, the only Rejects brought into the sample at all were the ones on which we reclassified. The goal of reweighting is to "weight up" the Accepts to stand for themselves and their like-scoring Rejects. Because the Accepts and

Rejects score similarly within each Score interval, it's reasonable to assume that their performances would be similar. Remember, the same assumptions were made in Parceling. The main differences are that parceling lacked the upfront reclassification and physically moved all the Rejects into the Accept group as either Goods or Bads. Augmentation merely weights the existing Accepts up. Below is an example of Score by Accept (after some data smoothing), performed with PROC TABULATE:

SCORE	REJECT #	ACCEPT #	TOTAL #	REJECT %	ACCEPT %	REJWGT
LOW-499	3165	3482	6647	47.62%	52.38%	1.9091
500-549	285	557	842	33.84%	66.16%	1.5115
550-649	380	857	1237	30.72%	69.28%	1.4438
650-749	214	727	941	22.74%	77.26%	1.2941
750-899	190	1183	1373	13.84%	86.16%	1.1607
900-HIGH	0	940	940	0.00%	100.00%	1.0000

REJWGT is the weight that will be applied to the Accepts. For example, in the score range of 500 to 549, the 557 Accepts will be weighted by 1.5115, taking them to 842, the total number of observations in that score interval. In this way, they will represent themselves and the Rejects in that group.

The SAS code that creates these weights is as follows:

```
REJWGT = 1.0; /* Initialization of REJWGT */

IF (SCORE <= 499) THEN REJWGT = 1.9091;
  ELSE IF (500 <= SCORE <= 549) THEN REJWGT
    = 1.5115;

  more lines of code

  ELSE IF (SCORE >= 900) THEN
    REJWGT = 1.0000;

IF (REJECT1 = 1) THEN REJWGT = 1.0;
  /* To unweight Rejects not reclass.*/
  ELSE REJWGT=REJWGT;
```

Once reclassification and reweighting have been accomplished, it's important to review the "new" Bads (BAD1) and compare them to the original (BAD). The reason for reject inferencing is to augment the known data (Goods and Bads) with the unknown (Rejects). If very few Rejects have been either reclassified or reweighted, the sample will still be primarily a known Bad-Good model. However, if too many Rejects have been brought over, the sample could be skewed into a Reject-Good model. This will underestimate the true Bads in the population, making the model less effective in identifying the Bads a creditor should not have

approved and the Rejects that he should have approved. Once a “correct” percentage of true Bads to Reclassified Bads has been determined, the weighting of the sample can be altered to adjust.

```
IF (REJECT = 1) & (DERPR >= 1) THEN
REJWGT=REJWGT*(some value);
ELSE REJWGT=REJWGT;
```

There are multiple statistical methods of calculating the correct ratio of known Bads to reclassified Bads. Industry practice usually falls into the range of 2:1 to 3:1 known Bads to reclassified Bads.

Weaknesses with Augmentation include the relative complexity of the calculations compared to Parceling and the larger number of assumptions about the data that are required.

## CONCLUSION

Because of the varying assumptions each of the above methodologies require, an apples-to-apples comparison is not possible. However, certain conclusions can be made about the relative merits of each. Cohort Performance works well when the industry data is clean and audited. Also, the creditor must make sure that his portfolio corresponds to the industry subset being compared. The lower the credit quality a lender is willing to accept, the harder it will be to get meaningful data on his Rejects, as fewer of these Rejects were able to obtain credit elsewhere.

Parceling is a quick, inexpensive, and relatively simple method of reject inferencing that requires no outside data sources. Heavy reliance on a single score on which to parcel makes using a stable score more important than ever. Also, portfolios with low delinquency rates allow for a limited number of

Rejects to be made Bads, making the results more of a known Good-Bad model than might be desired.

Augmentation is also quick, inexpensive, and requires no outside data sources, but it is a relatively complex procedure. These complexities lead to more assumptions and the possibility of more errors. However, the great flexibility Augmentation has in regards to both reclassification and reweighting, allowing for creditors of all types and credit quality to use the method equally well and under any conditions, makes Augmentation the best of the three.

## REFERENCES

SAS Institute, Inc. (1990), *SAS Guide to TABULATE Processing, Second Edition*, Cary, NC: SAS Institute, Inc.

## ACKNOWLEDGEMENTS

SAS is a registered trademark or trademark of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

## AUTHOR'S ADDRESS

The author may be contacted at:

Gregg Weldon  
Sigma Analytics & Consulting, Inc.  
1853 Peeler Road  
Suite D  
Atlanta, GA 30338

(770)804-1088  
gregg.weldon@sigmaanalytics.com