Dynamic Predictive Credit Scoring

with Multi-dimensional Analysis of Nearest Neighbors

Craig M. Allen, Ph.D. and Cherry Allen-Driscoll[1]

This paper discusses a database lookup methodology as an alternative to traditional credit- and behavior-scoring models. In this methodology, a statistically significant sample of "nearest neighbors" is drawn from a historical database to predict the performance of a subject loan / borrower / collateral unit. The similarity of the nearest neighbors is determined by utilizing a weighted distance measure in the principal component space of the historical loan / borrower / collateral / performance universe. The principal component rotations and the basis for determining weights in the distance space are re-calculated frequently, according to the volatility of the underlying data. This methodology has been applied in the automobile-loan sector with high degrees of success. It is argued that given the current chaos in the world credit markets, rapid modifications to predictive models as provided by database lookups are essential and provide a much more sound foundation for determining borrower / loan suitability than that provided by traditional credit scores or borrower score cards.

Most of the common methods used to predict loan performance behavior (i.e., "credit scores") require an initial segmentation of a given set of data. In the simplest case, this is a division of past borrowers into population segments of "those that defaulted on their payments" and "those that did not default." In this simple case, a discriminant function is usually developed using statistical methods that best separate these different populations. Recent innovations in credit scoring (e.g., the "Vantage Score" promoted by the three major credit reporting companies, see "Vantage Score" 2009) have proven that the market now realizes a simple division of historic populations into two broad segments is an over-simplification of the problem. These new models drill down into more granular subpopulations and within those subpopulations study a binary behavior, such as default or no-default. Thus, the industry practice has moved from a presumed homogeneous population to one of distinct subpopulations, each with its own discriminant function. Such a second generation predictive model creates a "credit score" as proxy for default likelihood, the scalar nature of which distracts from the fact that it is a synthesis of discontinuous sub-regions of the population space which were isolated at an arbitrary point in time.

The current paper argues that there is a better alternative. The current state of the art in database access techniques allows for a more robust, data driven approach to behavior prediction. Instead of relying upon *ad hoc* arbitrary segmentations of a borrower population and forcing piece-wise linear discriminant functions to give a scalar proxy for default likelihood, current data aggregation and access techniques allow a dynamic clustering approach, based on live data feeds and the most recent performance information to pull similar borrowers together into cohesive groups called herein "neighborhoods." These neighborhoods of similar borrowers can then be analyzed using a variety of

[1] Craig M. Allen, Ph.D. and Cherry Allen-Driscoll are currently associated with Delphi Analytics, LLC, a U.S. based boutique "quant shop" that focuses on loan portfolio analysis.

techniques to predict expected behavior. This approach, which has been called "Multi-dimensional Analysis of Nearest Neighbors" (Allen, 1999), provides a more dynamic set of behavior predictors than traditional approaches. These clustering-based techniques have sometimes been discussed based upon the methodology used to aggregate borrowers into groups. The MANN approach focuses on the "analysis" of clusters of borrowers rather than on extracting a single number representing some behavior's likelihood.

A focus on the aggregation technique will generally miss the major benefits of adopting a data-centric approach. As such, this paper is focused on the benefits of adopting the state of the art data-access methodologies (i.e., MANN). Such a data driven approach provides a multitude of benefits which better meet the broader objectives of the credit analysis, which include 1) a non-arbitrary segmentation of the borrower population into smaller clusters for analysis, 2) the measurement of multiple behavior measures, not just the binary division of a populations (or sub-population) into two, and 3) the establishment of a truly dynamic predictive framework that easily adapts itself to volatile economic and behavioral environments. **We argue that the rapid changes in global economic conditions, the extreme variations in employment rates and collateral values and the severe contraction of sources of family liquidity all demand a more dynamic process of credit analysis such as the dynamic segmentation available through MANN methods.**

Caprice in Predicting Loan Performance

As we descend deeper and deeper into economic chaos there is growing ridicule heaped on traditional "credit scoring" techniques. The abject failure of previous approaches begs us to focus on the most basic assumptions of previous modeling exercises.

At one level, the most damaging and arbitrary of all assumptions in lending (or investing) has been that borrower credit is a sufficient indicator of loan performance. As we now most painfully know, the value of a loan is very much a function of not just borrower credit, but of collateral value, of originator diligence or fraud, of loan parameters and of employment rates and other local economic indicators. Credit scores are interesting – but at best they describe only part of a picture, and perhaps not even the most important part, in predicting loan performance.

While "credit" might be a useful construct in predicting loan performance, we know that it cannot be examined to the exclusion of other predictors. One popular solution is to assign various risk multipliers to credit scores, market sectors, LTV, and increasingly long lists of granular categories and sub-segments[2][3]. However, such arbitrary weightings and subjective divisions and subdivisions of the portfolio into subpopulations actually constitute a step backward. We argue that "the numbers" should drive the modeling, not subjectivity.

[2] http://www.fitchratings.com/creditdesk/reports/report_frame.cfm?rpt_id=437248&sector_flag=3&marketsector=2&detail=

[3] http://dbrs.com/research/227912/rating-u-s-residential-mortgage-backed-securities-transactions.pdf

Among the arbitrary assumptions in credit modeling, it is well known to be an erroneous assumption that the progression from "good credit" to "bad credit" is in any way a linear function. It is only an artifact of early statistical discriminant techniques that we even think that a single number like a credit score has any bearing on loan performance. It has proven to be a simple and generally useful tool, as high scores are usually indicative of a borrower belonging to a sub-population that generally pays their bills – and low scores reflect the contrary. However, it is difficult to even imagine the population of borrowers measured by credit scores is monotonic in *any* predictive sense in *any* attribute dimension. It is almost the equivalent of suggesting real estate values are monotonically increasing as one travels in a line from New Orleans to New York City. Perhaps this is generally true, but only if one ignores the structure of neighborhoods and clusters of value of all other peaks and tro along the way.

With borrower credit, neighborhoods of exception and discontinuity have increasingly crept into credit scoring models. The following figure (Figure 1) shows the "traditional" segmentation approach (see *Vantage Score*, 2009) in accounting for these discontinuities. Here, an attribute, is used to sub-divide the population into smaller subgroups that are more homogeneous. It is these smaller homogeneous sub-groups that are then subjected to the linearization that fails on the populations as a whole.

Figure 1: Traditional Segmentation Scheme[4]

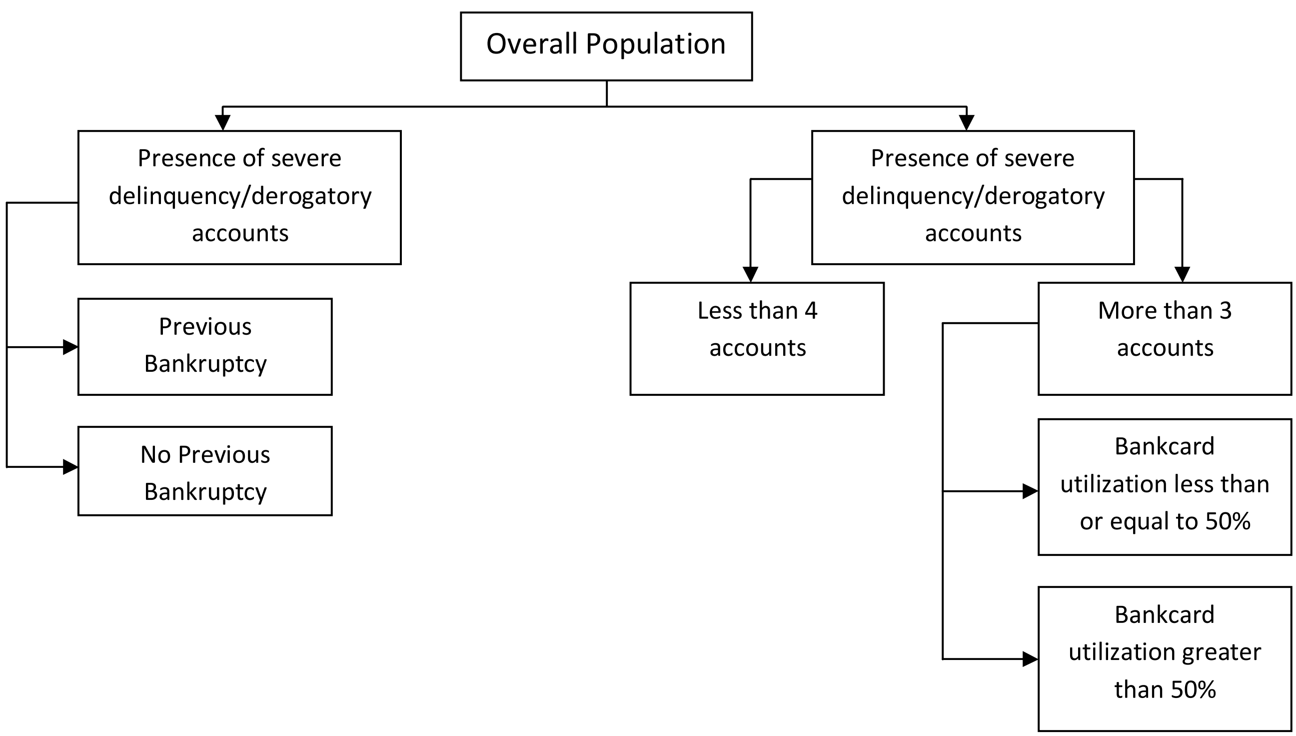

Figure 1: Illustration of traditional segmentation **(Vantage Score, 2006)**

Yes, this segmentation helps make the linear modeling approach more accurate – but only at the cost of introducing somewhat arbitrary segmentation rules. Why is bankcard utilization of 50% a meaningful dividing line within the populations? Why not at 37.25%?

4 (Vantage Score, 2006)

The process of identifying more granular subpopulations has gained momentum in credit scoring. William M. Makuch, in the Handbook of Credit Scoring, describes how more recent approaches to segmentation seek to group like populations, more than separate the goods from the bads. From these like populations, a different score card can be built for each population (Makuch, 2001). This concept is illustrated in the "Vantage Score". We ask, why is there such spurious divisions of a population? The answer is simply that subdivisions make the modeling work better. The arbitrary segmentation is justified by superior results.

Figure 2: Second Generation Segmentation Scheme[5]

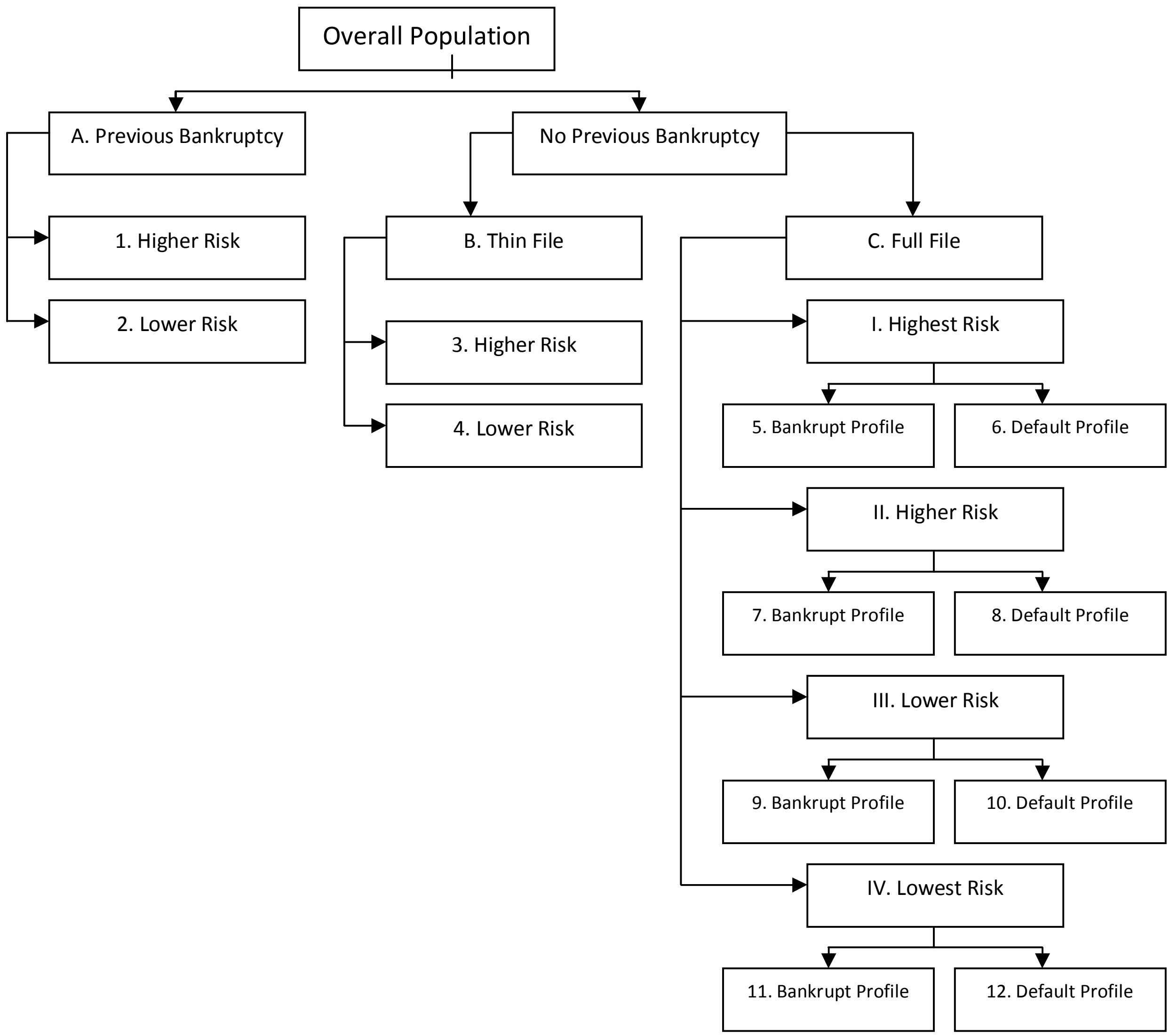

Figure 2: Illustration of second generation segmentation **(Vantage Score, 2006)**

However, the segmentation strategy illustrated above, does not accept feedback on its continued significance and represents an a priori setting that would require great efforts to modify. There is a price to be paid for allowing such caprice to enter into credit scoring systems used by regulated entities

5 (Vantage Score, 2006)

in choosing to which customers they grant or deny a loan. Even if such segmentation proves superior in predicting performance, we argue that subjectively determined segmentation is unacceptable. The institutional coordination required to make a subjective change in the applied segmentation is enormous. This is not simply pursuing a better model fit, depending on the institution, the segmentation strategies are frequently the result of an institution's operational process and less driven by the scoring mechanic itself. The first generation scoring models assumed homogeneous populations, the second generation models achieved better results through segmentation, but the act of segmentation is not dynamic, nor does it easily occur on its own. Instead, institutional decisions are required to initiate the process of pursuing segmentation strategies in an effort to improve the scoring results of the product. Such an approach becomes increasingly complex and expensive to monitor if and when changes in behavior warrant migrations of those original segmentations. In an objective credit model, as typified by a MANN approach, there would be no human element deciding which features need to be considered in segmenting a borrower space.

Limited Objectives of First and Second Generation Methods

Taking a step back from the details of the popular techniques used to derive credit scores, Eisenbeis pointed out that there are few who have devoted specific attention to the underlying objectives which should be achieved when formulating or selecting a model. He also reminds us that the objective of any credit analysis should be to maximize the expected net present value of a loan (Eisenbeis, 2004). As the capital markets evolve, primary loan originators increasingly seek additional balance sheet capacity through secondary market sales of whole loan or structured product transactions. As such, in addition to the credit quality of the borrower, the objective of modern credit analysis must incorporate quality measures of originations, servicing and collections, potential for fraud, and the value of any underling collateral. Creditors are constantly faced with the decisions of deciding how much information to collect before making a lending decision. As financial markets evolve, both the needs and opportunities to collect information are growing exponentially.

To some extent limited perspectives on the objectives of credit analytics have been illustrated recently by severely underestimated defaults in the US residential mortgage market. Many market players failed to identify deteriorating underwriting standards which led them to significantly underestimate defaults (Jenkinson, 2008). Prior to the housing/credit bubble, many credit ratings were calculated based on limited credit criteria that had been generally successful at categorizing credit risks. The over-reliance on the traditionally predictive categories which did not thoroughly incorporate origination quality coupled with stale organizational process fueled a failure to recognize that the world had changed. The traditional methods of categorizing and assessing credit quality were no longer relevant. As a result, the simple acts of verifying application data such as employment, income and assets are becoming increasingly significant, as secondary lenders recognize the potential credit risks embedded in the origination process.[6]

[6] For example, the American Securitization Forum has proposed a range of new fields for securitized mortgage loans which represent various levels of Income, Employment and Asset verification.

The MANN Method

Instead of designing segments by static and predetermined categories, MANN methodologies rely on a *K* nearest neighbor approach to identify similar peer groups (neighborhoods) from a multi dimensional information rich perspective. New borrowers are matched to the neighborhood where they belong and then the behaviors of concern are projected based on what has occurred historically within each group. The nearest neighbors approach is well recognized as one of the simplest approaches to obtaining a nonparametric classification (Arnaud de Servigny, 2004). The technique has furthered automated decision making in a wide range of sciences and various studies have demonstrated that non-parametric methods result in higher classification accuracies than parametric methods (M. P. Sampat, 2004).

To further reinforce dynamic segmentation, principal component analysis (or "PCA") defines the most important dimensions of space used to establish similarity measures among borrowers. PCA allows the segmentation method to remain dynamic, as it does not assume we know which variables or how many are important to measure. This approach consistently identifies the most meaningful basis to re-express a data set, filter out the noise and reveal the hidden structure of the data. PCA helps to simplify any number of potential variables while ensuring the measured data does not become unwieldy, clouded or redundant. The approach ensures each of the components is independent of every other principal component, and ensures the distance measures between points plotted in this principal component space are well behaved mathematically.

Figure 3: Nearest Neighbor Clusters

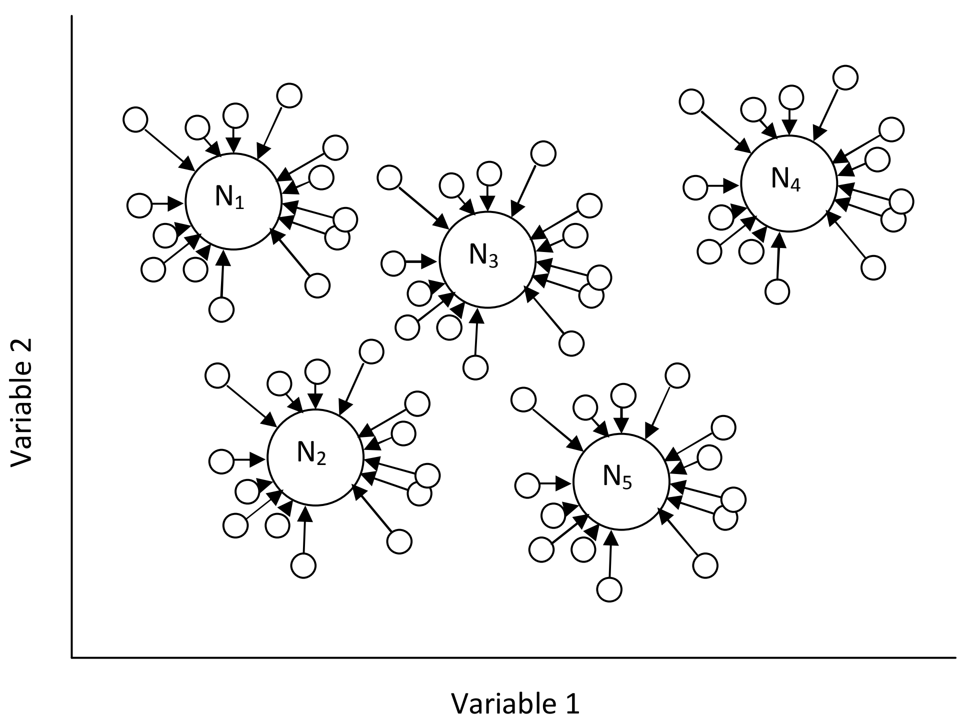

Figure 3: Illustration of nearest neighbor clusters.

The principal component rotations and the basis for determining weights in the distance space are re-calculated frequently, according to the volatility of the underlying data. As behaviors change and new information is acquired, neighborhoods and clusters are free to migrate as the information dictates. Such an approach broadens an understanding of behavior beyond the predetermined characteristic segmentations required by traditional methods.

Fluid Operations

A data-centric clustering approach meets a broad range of objectives for the credit analyst. It is capable of measuring a wide range of behaviors, often very complex behaviors, which integrate more directly into profitability measures, and it provides a framework for dynamic clustering, dynamic behavioral forecasts and consistent surveillance.

Behavior Forecasting - A More Flexible Solution

Using neighborhood methods, virtually any measure of interest can be monitored and hopefully predicted for a group of borrowers. For each neighborhood the future behavior of each loan is predicted by the history of each of the loan's peers within that neighborhood. Instead of assigning an intermediate credit score, a direct monitoring and assessment approach provides flexibility and a greater understanding of the real underlying characteristics of each cluster of borrowers. This ability to assess each measure of interest separately and to do so independently of any predetermined assumptions on segmentation gives the MANN methods extraordinary versatility. Various borrower behaviors in addition to expected default rates (such as prepayments, delinquencies, collateral damage, and costs of servicing) can be measured and predicted. Operational efficiencies are achieved as other measurable behaviors are monitored, measured and forecast with the same database and technology.

Lending decisions are made on loans that are expected to generate cash flows over long periods of time. A loan generates flows until it is either paid in full or defaults – and then upon default there is potential for recoveries through secured collateral or continued collection efforts. The value of the loan is not answered by a binary good or bad, value is determined by the length of time the loan is current (Eisenbeis, 2004). Scalar credit scoring models offer simple, easily understood results indicating probabilities of default; however, they avoid the problem of distributing those probabilities over time. The flexibility provided by MANN methods can more easily associate timing expectations with behavioral probabilities. The timing forecasts of expected behaviors translate directly into actual cash flow projections.

Dynamic Forecasts & Consistent Surveillance

One of the benefits of having a rich mathematical basis for determining borrower similarity is that extensions to the categorization model are easily made and understood. As a portfolio seasons, the MANN method incorporates updated and current payment behaviors to refine the accuracy of the future expectations. Access to a database of current attributes and/or monthly behaviors ensures that on each database access (or refresh), the borrower is compared to the currently most similar cohorts, and that the behaviors of those cohorts at that moment in time are those predicting the behavior of the borrower being referenced. Instead of stale data being locked into an immutable credit score, each access provides a sort of "recalibration" of the model.

As a loan seasons, its performance may provide loan grouping information that is useful in predicting future credit performance. The behavioral measures are applied, just as underwriting variables are used, in appraising the similarity of loans. With these additional similarity measures, the historical performance of each loan neighborhood (from that point forward in time) can be assessed. Different default rates and delinquency rates are often observed in neighborhoods defined, in part, by these early behavior measures. It is the flexibility of this data-mining approach that allows each defining characteristic of a loan to be used, as might be appropriate, in obtaining increasingly better estimate of expected borrower performance.

Loans, for example, with a payment lapse of 60 days or more within the first N months of their life may be tagged and event frequencies counted. Each occurrence is identified as either a lapse with recovery or a lapse with no recovery. The number and category of each type of delinquency aids in grouping seasoned loans and enhances predictive ability for the remainder of their life. For example, 24-month-old loans with two or more 60-day lapse-with-no-recovery delinquencies typically have a higher expected default rate than those with no 60-day delinquencies during their first 24 months, and so forth[7].

The recalibration of forecasts ensures that all available information is used as a foundation to refine predictive accuracy on a real time basis. This ensures a consistent methodology based on current and updated database fields. The consistency from initial scoring through the seasoned life provides operational simplifications that are not possible with many other methods. Expected future behaviors do not become stale at origination only to be refreshed when new assessments are required or economically desirable. Measurable and consistent processes can be updated to ensure these critical valuations are reflecting realistic future expectations.

The benefit of this approach is the unlocking of the richness of information contained in historical payment files. When properly set up and repeatedly updated with each month's history added, this approach leads to a dynamic and self-correcting estimation system for loan performance. The calculations are intense, but the value of increased predictive ability usually outweighs the cost of the computer and analysis.

Summary and Conclusions

The benefits of adopting a data-centric methodology for the non-arbitrary segmentation of the borrower population into smaller clusters are many. This general approach meets a broad range of objectives for the credit analyst and for the analyst more broadly charged with identifying "profitable" lending opportunities. Such a framework is capable of measuring a wide range of behaviors, it offers

7 Delphi Analytics

more complex types of analysis which integrate well into profitability measures, and it provides a framework that offers dynamic clustering, dynamic behavioral forecasts and consistent surveillance.

Bibliography

Allen, C. M. (1999). Credit Scoring and Risk-Adjusted Pricing: A Review of Techniques. In F. J. Fabozzi, *Subprime Consumer Lending* (p. 38). New Hope: Frank J. Fabozzi Associates.

American Securitization Forum. (2009). *ASF RMBS DISCLOSURE AND REPORTING PACKAGES Final Release.* http://www.americansecuritization.com/uploadedFiles/ASF_Project_RESTART_Final_Release_7_15_09.pdf.

DBRS. (2009). *Rating U.S. Residential Mortgage-Backed Securities.* New York: DBRS.

de Servigny, A., & Renault, O. (2004). *Measuring and Managing Credit Risk: Quantitative Approaches for Default Risk/Data Analysis and Models for Loss Distrubutions/Unique Strategies for Bank Capital Allocation and Securitizat.* McGraw-Hill Professional.

Eisenbeis, R. (2004). Problems in Applying Discriminant Analysis. In D. B. L. C. Thomas, *Readings in credit scoring: foundations, developments, and aims.* Oxford University Press.

FitchRatings. (2009). *ResiLogic™: U.S. Residential Mortgage Loss Model Criteria.* New York: FitchRatings.

Jenkinson, N. (2008). *CGFS Papers No 32: Ratings in structured Finance: What went wrong and what can be done to address shrotcommings?* Committee on the Global Financial System.

Makuch, W. M. (2001). The Basics of a Better Application Score. In E. Mays, *Handbook of Credit Scoring.* Lessons Professional Publishing.

Sampat, M. P., Bovik, A. C., & Aggarwal, J. K. (2004). Supervised Parametric and Non-Parametric Classi¯cation of Chromosome Images. *Elsevier Science* , 21.

Vantage Score. (2006, May). *Segmentation for Credit Based Delinquency Models White Paper*. Retrieved July 21, 2009, from Vantagescore: http://www.vantagescore.com/docs/segmentation.pdf