WHITE PAPER

The case for multiple approaches to retail credit portfolio analysis

Roger M. Stein*

First Draft: December 1, 2010
This Draft: May 18, 2011

Abstract
Practitioners apply various methods of portfolio analysis to the evaluations of the credit risk of retail debt. One way to characterize these approaches is to differentiate them along two dimensions: the *unit* of analysis (loan-level vs. pool-level) and the *state-space* of the analysis (single path vs. full distribution of paths). We provide a heuristic framework for evaluating these tradeoffs and highlight instances where one framework may be preferred to another. It is often the case that additional factors, besides quantitative ones, bear on the analytic choice. These factors may relate to organizational or domain-specific constraints or preferences. Combining the various forms of analysis may provide a more holistic view of both the models and the reference portfolios.

1 Introduction[1]

In this non-technical white paper we discuss considerations that retail risk managers face in combining analytic approaches with organizational applications.

While much has been written on the quantitative modeling of retail credit risk, the literature provides little guidance on what types of models are best suited to which types of real-world applications. The vast majority of discussions among academics about the comparative merits of various analytic approaches center on the relative quantitative efficacy and

* Moody's Research Labs, New York, NY. email: roger.stein@moodys.com

[1] I am grateful to Ashish Das, Rob Fauber, Felipe Jordao and Jody Rasch for their comments and suggestions. Paul Getman provided a number of very helpful suggestions on presentation and was also instrumental in working through the use cases. I am particularly indebted to Mark Zandi for extensive conversations on earlier drafts of this document and for his detailed suggestions on improving the presentation and content.

performance of different techniques or models. However, in applied settings, it often happens that quantitative properties are only one of the objectives in performing an analysis.

In implementing retail risk systems, we have found it helpful to think about not only the analytic properties of the tools, but also their practical business and organizational implications. We examine some of these issues and make a number of stylized recommendations for sample applications. To anticipate our discussion: for forecasting and risk applications in retail credit, we generally recommend *combining various analytic approaches*. However, these different approaches should be used as *complements* to each other rather than as redundant measures. The combined approaches tend to form a more complete picture of the risk in key market segments as they may relate to a specific portfolio. While our view is that a multifaceted approach is most informative, we do have clear recommendations regarding which components of the risk picture are better addressed with which types of models.

Note that since this is not a research paper, we do not present new empirical or analytical results. Rather in this white paper we discuss a number of workaday issues that users must address in selecting modeling technology. In our experience, users often struggle in determining how best to address their analytic needs in the face of a number of potentially attractive solutions. Our goal is to facilitate this decision process. We make liberal use of shorthand in our descriptions and discussions when we judge that our meaning will be obvious to practitioners.

The remainder of this paper is organized as follows: Section 2 discusses a simple conceptual taxonomy of various approaches to retail credit risk. Since quantitative performance is often the predominant focus of many discussions of model fitness, Section 3 discusses some of the quantitative properties of different approaches. Section 4 provides a model "fitness" framework that contemplates both quantitative and organizational/business considerations in model selection and evaluates the methods outlined in the taxonomy using this framework. Section 5 provides a series of example target applications. The goal is to demonstrate the use of the framework as a tool to evaluate how well different approaches meet the analytic and business requirements. Section 6 discusses why a user could find it useful to combine multiple approaches in performing risk analysis.

2 A framework for categorizing analytic approaches to retail portfolio credit risk

2.1 Preliminary definitions

Before we begin the discussion of the fitness of different approaches to modeling retail credit risk, it is useful to define terms. While meanings of these terms may vary across researchers, for purposes of this article, we adopt the following definitions:

Aggregate analysis or *aggregate-level analysis* is a term for a model or models estimated using data that is constructed by combining individual observations into some form of cohort. This would include, most obviously, the pooling of an entire portfolio into a single cohort or sub-cohorts, but it would also apply to the pooling of MSA-level delinquencies to form state-level delinquencies. Aggregate models take as input some vector of descriptive statistics for the cohort (e.g., mean LTV or mean FICO score) and use these to predict the corresponding aggregate measure of interest (e.g., the default rate for the cohort). Note that the construction of cohorts need not be done at the time of model construction. For example, an aggregate model may be built from the stratification of individual loans in a user's retail portfolio or it may be built from aggregate governmental statistics on regional default rates.

Loan-level analysis or *micro-level analysis* refers to a model or models estimated using data for individual loans directly. If the measure of interest is an aggregate one, such as the expected loss for a portfolio, the aggregate measure is estimated by first estimating the unit-level analog (e.g., the loan-specific EL) and then combining these unit-level measures to form the aggregate measure. Loan-level models take as input some vector of descriptive statistics for the individual loan record (e.g., its LTV or FICO score) and use these to predict the loan-level behavior.

Single path analysis denotes an analysis done using a stand-alone economic or behavioral scenario. Such analysis is commonly used for stress testing. For example, an analysis may be done using a single economic forecast or using the macro-economic path from a particular historical period. However, the path need not refer only to macro-economic data. It may be described using, e.g., CPR and CDR timing curves which refer only to prepayment and default rates and timing, without necessarily requiring explanations of *why* the levels and timing take particular values. Also note that it is not necessarily the case that the term "path" refers to a single time series. In the case of a macro-economic path, for instance, a single path may comprise several economic time series (e.g., interest rates, GDP and unemployment) and, furthermore, these time series may be for a number of geographic regions.

In practice, several or even dozens of paths may be used for such analysis. However, the key feature of these paths is that they are not generated by a strictly statistical process. Rather they may be created by economists or traders based on subjective inputs or based on a mechanical generation of a grid of values for different factors (e.g., an analyst may construct a grid by creating 20 scenarios in which interest national rates move in 25bps increments over a range of +/- 2.5% from the current. More involved grids can be constructed using additional factors.) This type of sensitivity analysis does not produce probability distributions, however.

Distribution of paths or *simulation analysis* refers to methods that algorithmically generate a large number of *single paths* and do so with explicit reference to the joint distributions of the factors in the paths. Thus, *more likely* scenarios are represented more often in the distribution than less likely ones. (Note the difference between this and the grid approach discussed above.) Because the frequency of paths correspond to their natural probability of occurrence (under the model), the distribution of paths may, in turn, be used to generate a loss distribution or distribution of cashflows that has some of the required properties for risk management and pricing. For example, these distributions can be transformed into loss or cashflow distributions for which explicit loss or payout probabilities can be calculated.

Though, for low-dimensional processes, it is possible to generate distributions of paths analytically, most practical applications require Monte-Carlo simulation. For path dependent processes, it can be particularly difficult to generate path distributions in analytic form. For this reason, we informally refer to these approaches as *simulation* approaches, though simulation is only one of the approaches to accomplishing this numerically.

Next, we outline a taxonomy for describing models. We find the taxonomy particularly useful in clarifying for business people the different configurations that can be accommodated by different combinations of modeling approaches.

2.2 A simple taxonomy for retail credit risk models

One way to provide a taxonomy of different approaches to retail credit portfolio risk is to differentiate the approaches along two analytic dimensions: the *unit* of analysis (loan-level vs. pool-level) and the *state-space* of the analysis (single path vs. full distribution of paths).

Figure 1, below provides a schematic of this interpretation. In the figure, the *x*-axis (horizontal) differentiates approaches by the unit of analysis, i.e., whether the analysis is done at the aggregate- (portfolio- or pool-) level or at the micro- (loan-) level. The *y*-axis (vertical) differentiates the approaches by the *state space* used to span the behavior of the portfolio, i.e., either a single paths or a full distribution of paths. Here, "single-path" may be generalized to a number of paths (typically few in number), but retains the notion that the path is generated by some method that does not explicitly rely only on the joint probability distribution of the underlying factors to produce a full distribution of economic outcomes. Analyzing one or a few macro-economic stress scenarios on a mortgage portfolio would be an example, as would running stressed CPR and CDR curves.

Figure 1 A taxonomy of retail credit risk analytic approaches

Unit of analysis

State Space	Loan-level	Aggregate-level
Single path	EXAMPLE: • Run single economic path over each loan in sample using individual loan characteristics • Produce loan-level cashflows, CDR & CPR curves, etc. • Aggregate loan-level output if necessary	EXAMPLE: • Run single economic path over aggregate sample, using aggregate information on loan characteristics • Produce aggregate-level CDR & CPR curves, loss curves, etc.
Distribution of paths	EXAMPLE: • Run *N* economic paths (*N* >> *1*) over each loan in sample using individual loan characteristics • Produce loan-level cashflows, CDR & CPR curves, etc. for each path • Aggregate loan-level output if necessary • Produce full distribution of losses, cashflows, etc.	EXAMPLE: • Run *N* economic paths (*N* >> *1*) over aggregate summaries of loan characteristics in CA • Produce distribution of delinquency rates in CA

In the next section, we discuss briefly some *quantitative* properties of the different approaches outlined in Figure 1 and then, in Section 4, go on to extend our analysis to what we believe is a more practically useful set of dimensions. These include not only measures such as precision, but also a number of other key organizational and business factors that come up in implementation. These other factors can supersede one-dimensional measures of quantitative performance in many settings. In Section 5, we use the full business frameworks to evaluate some sample applications.

3 Quantitative properties of different modeling approaches

We discuss the most common criterion used to differentiate analytic approaches: the *quantitative precision* of the tools. We do this in a stylized manner recognizing that each implementation will have different characteristics. Despite this, we focus on the most salient features of each approach in our comparisons.

3.1 Comparing single-path analysis to analysis of the full economic distribution

The quantitative distinctions between single-path and full simulation analysis typically come down to two features: the ability to assign meaningful probabilities to model outputs (which favors simulation approaches) and the ability to evaluate outcomes that may be outside of the natural distribution of economic paths (which favors scenario analysis and stress testing). We judge both of these to add value to prudent risk management.

With respect to the discussion of stress testing vs. full-blown simulated distribution analysis, we observe that, in general, if the objective is a quantitative one, distributional and simulation tools provide a richer probabilistic infrastructure for risk management. This observation is driven by the great difficulty in assigning probabilities to losses in a portfolio, even when the probability of a particular economic path is specified. This is true for both, "baseline" or "average" economic scenarios and for scenarios designed to stress a portfolio and thereby explore the tail of the portfolio's loss distribution.[2] This implies that neither the mean (expected) loss nor the tail-losses (economic capital) may be inferred from scenario analysis, even if point probabilities are assigned to the macro-economic scenarios, themselves.

However, this leaves open the question of how to think about specific economic cases that a regulator may require or for which an economist may have a strong view. This type of intuition based analysis provides a key component of the risk profile of a portfolio. It also provides a mechanism to assess the face-validity of a simulation output. Stress testing with macro-economic scenarios that are well understood provides a good reality check and validation for a model and, in some cases, for a portfolio's behavior.

More importantly, simulation-based approaches, by construction, provide no natural mechanisms to contemplate or evaluate states of the world that are outside of the parameters of the simulation itself. The great strength of stress-testing with single scenarios is that it permits users to define arbitrary macro-economic or asset-loss scenarios and to do so in

[2] To see this, consider first the mean case in which we have a baseline or "average" economic scenario. It turns out that the average losses for the portfolio will not be the same as the losses under the average scenario: Consider the stylized example of a single loan with an LTV of 80%. In this example, assume further that the loan only defaults when its LTV rises above 80%.

Imagine that in the average economic scenario, asset prices are flat, so LTV does not rise above 80%. Thus, under the average scenario, the loan does not default and has a default rate of 0%. However, consider what happens in bad states of the economy: home prices fall and the loan defaults. If this happens in 1% of all cases, then the mean default rate across all scenarios will be 1%.

But recall that in the case of the average path, we observed a default rate of zero. This is in contrast to full distribution approach, which suggests a default rate of 1%. This happens even when the mean of the simulated economic paths is *exactly* the same as the mean economic scenario used in the simulation. Similar, though more mathematically involved, arguments demonstrate analogous results for losses under many scenarios, even when the probability is explicitly assigned to the macro-economic scenario. See Chinchalkar and Stein, (2010, Appendix A) for a more detailed analysis.

a concrete and economically intuitive fashion. Stress-testing allows users to gain insight into how their portfolios might behave (under their model) in states of the world that may never have been contemplated by a particular model and for which there is no observable data.

3.2 Comparing aggregate analysis to loan-level analysis

Quantitative differences between loan-level and aggregate-level analyses center primarily on the different approaches' abilities to deal with portfolio heterogeneity and the non-linearity of relationships between risk and asset characteristics. Most loan portfolios exhibit heterogeneity (while the degree of non-linearity of risk varies by product type). Our observation is that, in general, loan-level analysis provides a more detailed decomposition of the risks in an individual portfolio. This detail can be important in some cases, and the differences between loan- and aggregate-level analyses can be particularly large for heterogeneous portfolios.

In our usage, *heterogeneity* implies that there are clusters of loans with similar attributes, but that loan attributes are not distributed evenly across all loans. This often leads to risk layering within the portfolio. The challenge is that aggregate portfolio summaries (e.g., average FICO) do not capture these interactions. In some asset classes (e.g., mortgages) where the asset risk behavior is also non-linear, this problem can become more even pronounced.

For example, we might imagine that two portfolios have identical average LTV and FICO statistics. However, in one portfolio, FICO and LTV are independent, while in the other the low FICO scores tend to be associated with the high LTV loans. The risk profiles of these two portfolios will be markedly different in many cases even though the two portfolios will look identical based on aggregate statistics. It is often difficult to determine whether such heterogeneity increases or decreases risk in a particular setting.[3] Of course, expanding the analysis from two factors to the many factors typically associated with credit risk, increases the complexity of the analysis.

Thus, in general, provided loan-level data and models are available, *unless heterogeneity can be ruled out as a complicating factor,* loan-level approaches are recommended from a quantitative standpoint. This is consistent with the recent econometrics literature and with some of our own results (see Chinchalkar and Stein, 2010, and references therein, for a detailed discussion of this phenomenon and for examples in the domain of residential mortgages).[4]

---

[3] See Chinchalkar and Stein (2010) for detailed empirical examples.

[4] In principle, with sufficient data, it might be possible to capture the appropriate levels of granularity within a portfolio with sub-portfolio-level summaries. One possible approach is to exhaustively describe all conditional relationships within the data. For example, if there were k factors in the model (FICO, LTV, MSA, loan type, etc.) each with m levels, we would create $m \times k$ cohorts in order to capture the joint distributions of these factors. However, adequately including all such interactions would require so much data as to be impractical

However, though loan-level models are well suited to uncovering the risk structure of a specific portfolio and to highlighting how macro-economic factors affect it, aggregate models are typically better from a practical perspective at integrating information from *across* a market.

As an example, an aggregate model provides a more natural practical framework to evaluate broad shifts in regional or vintage credit quality. By integrating broad patterns in the behaviors of different asset classes in different geographic locations, aggregate models can leverage this far-reaching data enabling a more holistic examination of a specific market segment or asset class evolution. This is particularly useful for analysis of markets or assets that are not currently in the portfolio. Loan level models, in contrast are typically more limited in their ability to do this due to their narrow focus on the impact of macro-economic factors on the specific loan portfolio.

4 A framework for selecting the fit between an analytic approach and a domain-specific application

We next discuss the *non-quantitative* differences between modeling approaches, which, in many settings, are even more important for the organization than an understanding of the quantitative aspects of the model alone. Dhar and Stein (1997) introduced a framework (hereafter *DS*) for evaluating the fit of various analytic tools to specific organizational applications. We adapt this framework to the problem of retail portfolio credit risk, modifying and extending it where necessary.

The DS framework was originally developed for matching modeling technologies to decision problems by evaluating the fit of each proposed solution to a particular decision problem with respect to:

- the dimensions of the most important *analytic and business* requirements (or constraints) of the application; and
- the inherent *strengths or weaknesses* of the proposed solution along these key dimensions.

In contrast to standard "specifications documents," the objective of DS analysis is to evaluate the *analytic* dimensions of a problem rather than generic IT requirements. For example,

---

in most cases. For example, for a heterogeneous portfolio containing just the four factors in parentheses, one could construct 4 coarse levels for FICO (50 point increments between 600 and 750); four levels of LTV (70<, 70-80, 80-90, >90+); two loan types. If one further identified only 10 MSAs in the portfolio, this would require $4 \times 4 \times 2 \times 10 = 480$ stratifications to be parameterized. These stratifications would not account for other features such as the underlying value of the collateral, the coupon of the loan, loan seasoning, loan size, delinquency status or underwriting vintage, all of are important for stress-testing and risk assessment and which would quickly expand the number of stratifications into the tens of thousands.

one dimension of the DS framework involves *scalability.* However, in the DS context, scalability has much less to do with database speed or memory size and far more to do with the ability to adapt a modeling approach to many new domains. Some applications (e.g., creating a consistent tool to assess risk across all geographic regions) would require high levels of scalability since the modeling approach would need to be scaled up to cover many different markets. Tools that require extensive calibration for each market might score lower on this dimension than would tools for which calibration can be done automatically.

The original DS framework contemplates 14 dimensions of fit between a proposed solution and the organizational and analytic needs and constraints of the application. A tool is deemed a "good" solution to a domain-specific problem if it satisfies the constraints of the domain without requiring substantial additional modifications to the tool. The table, below, outlines the various dimensions along which we will evaluate retail credit portfolio tools. We have somewhat extended and modified the DS framework to accommodate the nuances of credit portfolio analysis.

In the interest of brevity, we provide only a thumbnail description of each dimension as the business challenges and constraints will be familiar to most readers. Interested readers can find more detailed descriptions of many of the dimensions in Dhar and Stein (1997) [5] .

Table 1 Brief descriptions of some dimensions for assessing fit between model type and application

Dimension	**Description: The degree to which the modeling approach…**
Precision of predictions	… produces estimates that agree with the behavior of the assets or portfolio.
Probabilistic interpretation	… can assign meaningful probabilities to the model outputs[6]
Explainability	… can be described easily to users and management.
Asset class scalability	… it is easy to add new asset classes using the approach.
Geographic scalability	… can be extended to other the same asset class in other geographic regions.
Tolerance for sparse data	... can be used in settings with limited input data.
Tolerance for heterogeneity	… can reflect the risks of heterogeneous assets within and between portfolios.
Computation speed	… can be applied at high speed (e.g., on a trading desk).
Transparency	… permits users to ask and answer questions about the drivers of model output.
Granularity	… allows users to drill down into a portfolio to understand assets that are driving risk.
Comprehensiveness	… produces a reasonably complete description of the distribution of credit losses.

The remainder of this section evaluates the approaches in Figure 1 along in each quadrant of the dimensions of Table 1. We present this analysis in summary form as the assessments are generally obvious. In the discussion that follows we assume that all of the analytic options we evaluate in the rest of this paper involve choices between best-of-breed approaches. Thus, in our examples, choices come down to the native properties of each approach rather than the fact that one specific model was constructed better than another.

[5] For a more direct application in the context of corporate credit risk, see Kumar, Stein and Assersohn (2006).

[6] Note that this is different than assigning probabilities to, e.g., a particular economic scenario.

We also note that different implementations will have different values on certain dimensions, but that we have tried to capture the regularities across different approaches and to keep our assumptions about each model class as generic as possible. To provide an anchor for readers, we outline our assumptions for each of the approaches as follows:

Aggregate scenario analysis: We assume that the approach involves using aggregate portfolio statistics (e.g., mean borrower FICO score, mean LTV, etc.) to estimate aggregate portfolio behavior (e.g., default, prepayment, etc.), subject to a given macro-economic scenario or default rate scenario. This is the upper right quadrant of Figure 1.

Loan-level scenario analysis: We assume that the approach involves using loan level data (e.g., borrower FICO score, loan LTV, loan type, etc.) to estimate individual borrower behavior (e.g., default, prepayment, etc.), which is then aggregated to form an estimate of overall portfolio behavior, subject to a given macro-economic scenario**.** This is the upper left quadrant of Figure 1.

Loan level simulation analysis: We assume that the approach involves using loan level data (e.g., borrower FICO score, loan LTV, loan type, etc.) to estimate individual borrower behavior (e.g., default, prepayment, etc.), which is then aggregated to form an estimate of overall portfolio behavior, and that this is done individually for a very broad distribution (e.g., thousands of possible paths) of macro-economic states of the world. This is the lower left quadrant of Figure 1.

Table 2 provides a description, using a HIGH-MEDIUM-LOW scale of each approach along each of the dimensions. The dimensions have been constructed such that a value of HIGH, all else equal, is preferred to a value of LOW and that it is harder to satisfy a HIGH requirement on an application than it is to satisfy a LOW one.

Table 2 Assessment of different broad modeling approaches along different dimensions of Table 1

Dimension	Aggregate scenario analysis	Loan-level scenario analysis	Loan level simulation analysis
Precision	MODERATE-HIGH[7]	HIGH	HIGH
Probabilistic interpretation	LOW	LOW	HIGH
Explainability	HIGH	HIGH	MODERATE
Asset class scalability	MODERATE-HIGH[8]	MODERATE	MODERATE
Geographic scalability	MODERATE-HIGH[8]	MODERATE	MODERATE
Tolerance for sparse data	HIGH	LOW	LOW
Tolerance for heterogeneity	LOW	HIGH	HIGH
Computation speed	HIGH	HIGH	LOW-MODERATE[9]
Transparency	MODERATE-HIGH[10]	HIGH	MODERATE
Granularity	LOW	HIGH	HIGH
Comprehensiveness	LOW	LOW	HIGH

In the next section, we present nine examples of specific applications and how the framework can be applied to select a modeling approach that fits well in each case.

5 Some sample applications

We present examples of a number of typical risk management and forecasting applications. In each case, we show how the framework of Section 4 can be used to identify promising modeling approaches. In cases where one or two dimensions dominate (e.g., tolerance for sparse data) we show less binding dimensions in grey, rather than black. This does not necessarily imply that these dimensions are less important, but rather that in the absence of the black dimensions, the gray ones cannot be achieved. In each case, we provide a brief description of the application and a rationale for our assessment of the key dimensions. Then, we evaluate each technique based on the assessments of Table 2.

---

[7] Depends on the degree of heterogeneity of the asset class.

[8] Depends on availability of aggregate data on the asset class at the appropriate level aggregation (e.g., national- vs. city-level).

[9] Depends on implementation and portfolio size.

[10] Depends on degree of interaction among asset behaviors (e.g., prepayment and default)

To keep the examples concise, we present only the key dimensions that differentiate the approaches. For example, even if Transparency is an important dimension, if all approaches meet the Transparency criterion, we do not show it in the tables. A summary of all nine cases, and the implied modeling approaches, can be found in the Conclusion (Section 7). In addition, it sometimes happens that one dimension dominates all others (e.g., if loan-level data is not available, Tolerance for Sparse Data will be a "deal breaker" and dominate all other requirements.) In these cases, the dominant dimension or dimensions are presented in black while the others are presented in gray.

Importantly, we have chosen a variety of applications to demonstrate the flexibility of the framework and to show how different dimensions bind on modeling decisions in different contexts. However, our list is far from exhaustive. There are a fair number of applications that are not listed but for which we could make clear recommendations. There is nothing special about the cases we have chosen, beyond that they provide a canvas on which to demonstrate the framework. We expect that some readers will wish to create similar analyses for their own applications and to extend the framework as needed.

5.1 Regional macro-economic stress testing and forecasting for a single asset class (loan-level data not available)

Description: Forecast the behavior of a key single class in a geographic region based on stress scenarios. For instance, user may be interested in forecasting the level of delinquencies in California over the next two years. An example of such analysis is presented in Zandi, *et al.*, 2009.

Key dimensions: This application lies at the heart of macro-economic forecasting and policy analysis. It focuses on identifying broad patterns in the evolution of a key asset class, subject to a specific set of stress scenarios. Furthermore, this analysis must typically be done without detailed micro-data on each segment and, ideally, by incorporating feedback loops over time. *Importantly, because the approach relies on aggregate data series as inputs, we assume that predictions are better using an aggregate model.*

Dimension	REQUIREMENT	Ag. Scenario	LBL Scenario	LBL Simu
Precision	HIGH	✓		
Tolerance for sparse data	HIGH	✓		
Geographic scalability	Depends on application	✓	Depends on application	Depends on application

5.2 Stress testing all retail portfolios across bank within a common model (loan-level data or loan-level models not available/practical for asset classes)

Description: Use a common framework to apply stress test to all assets in a multinational bank's portfolio, including regions and asset classes for which there is no loan-level data.

Key dimensions: This application provides our first example of a case in which even the best solution may still not satisfy all constraints. Ideally a modeling approach would address issues of potential portfolio heterogeneity, however, the need for scalability along both asset classes and geographies dominates these concerns. By construction, using a common framework across all asset classes precludes using any of the LBL approaches, because in this example, for many asset classes loan-level data are not even available to run the models.

Dimension	REQUIREMENT	Ag. Scenario	LBL Scenario	LBL Simu
Precision	HIGH		✓	✓
Asset class scalability	HIGH	✓		
Geographic scalability	HIGH	✓		
Tolerance for sparse data	HIGH	✓		
Tolerance for heterogeneity	HIGH		✓	✓

5.3 National economic policy evaluation

Description: Evaluate the impact of several proposed economic policies (e.g., reduction in interest rates, provision to purchase distressed assets, etc.) on the future state of the economy. An example of such analysis is presented in Blinder and Zandi (2010).

Key dimensions: This application focuses on identifying broad patterns in the evolution of the economy across many sectors. As such, the ability to evaluate different asset classes in many regions of the country and to compare these to other aggregate measures that result from the policy analysis (e.g., local unemployment or home-price levels) is paramount. Ideally, the analysis would also incorporate feedback between macro-factors over time. Furthermore, this analysis must be done without detailed micro-data on each segment.

Dimension	REQUIREMENT	Ag. Scenario	LBL Scenario	LBL Simu
Asset class scalability	HIGH	✓		
Geographic scalability	HIGH	✓		
Tolerance for sparse data	HIGH	✓		

5.4 Capital threshold analysis for key asset classes (loan-level data or loan-level models not available)

Description: Determine the whether the amount of capital required to support a retail portfolio under one or more stress-scenarios for a key retail asset class. Do this without access to loan-level data.

Key dimensions: The user wishes to estimate how much of a capital cushion might be required to support the bank's portfolio but the lack of data prevents drilling down into the data or considering heterogeneity. The approach must, thus, be able to produce estimates without micro data.

Dimension	REQUIREMENT	Ag. Scenario	LBL Scenario	LBL Simu
Precision	HIGH	Depends on asset class	✓	✓
Probabilistic interpretation	HIGH			✓
Tolerance for sparse data	HIGH	✓		
Tolerance for heterogeneity	HIGH		✓	✓
Comprehensiveness	HIGH			✓

5.5 Capital allocation for key asset classes (loan-level data and models available)

Description: Determine model-based economic capital required to support the portfolio for a key retail asset class, given a bank risk tolerance. Measures may include VaR or Expected Shortfall (See Bohn and Stein (2009) for a discussion). Determine which assets are most risky with respect to economic capital required and which ones increase diversification.

Key dimensions: This application goes to the heart of credit risk management. The user wishes to determine how much of a capital cushion is required to support the bank's portfolio with some probability. This requires a full distribution of losses resulting in probabilistic interpretation. It also requires the user to be able to drill down into the portfolio to understand which segments are driving the bank's risk.

Dimension	REQUIREMENT	Ag. Scenario	LBL Scenario	LBL Simu
Precision	HIGH	Depends on asset class	✓	✓
Probabilistic interpretation	HIGH			✓
Tolerance for heterogeneity	HIGH		✓	✓
Granularity	HIGH		✓	✓
Comprehensiveness	HIGH			✓

5.6 Stress testing heterogeneous mortgage portfolio (loan-level data and models available)

Description: Determine losses on a portfolio of mortgages that are heterogeneous with respect to loan-type, coupon, credit-score, etc., using one or more stressed macro-economic scenarios.

Key dimensions: The user wishes to stress the portfolio under a variety of macro-economic scenarios. The portfolio is not homogeneous in that there is variability in loan-type, margin, etc.

Dimension	REQUIREMENT	Ag. Scenario	LBL Scenario	LBL Simu
Precision	HIGH		✓	✓
Granularity	HIGH		✓	✓
Tolerance for heterogeneity	HIGH		✓	✓

5.7 Investment and loan selection/origination

Description: Develop a pricing system to allow individual loans to be priced, or, conversely, to permit loan terms and coupon to be determined given the credit risk of the borrower. (Optionally, in a fully functional implementation, it would also require that the transfer price be calculated. This is discussed in Section 5.8.)

Key dimensions: This application requires that the model produce a loan specific price or permit users to determine the credit risk of the loan in order to set an appropriate loan coupon and loan terms. This requires that individual loans be examined probabilistically and that the various borrower characteristics be evaluated in the context of various loan types.

Dimension	REQUIREMENT	Ag. Scenario	LBL Scenario	LBL Simu
Precision	HIGH	Depends on asset class	✓	✓
Probabilistic interpretation	HIGH			✓
Tolerance for heterogeneity	HIGH		✓	✓
Granularity	HIGH		✓	✓
Comprehensiveness[11]	HIGH			✓

5.8 Transfer pricing

Description: Develop an internal measure of the risk taken by individual lending officers and traders such that they can be charged for the capital that must be allocated to the risks to which they expose the bank. The risk should be measured in terms of the change in economic capital that results from the loan or trade. This will be used as one means of aligning the incentives of managers and line staff. See Bohn and Stein (2009) for a discussion and examples.

Key dimensions: This application requires that the model produce a *portfolio referent* measure of capital required to support the new position or loan, given the current portfolio holdings. It requires an analysis of the correlation of the loan with the whole portfolio and that an explicit probability and level of loss be assigned to the loan. This necessitates a full distribution of losses with corresponding probabilistic interpretation. It also requires the user to be able to drill down into the portfolio to in order to calculate the change in capital introduced by the new loan or position.

[11] Requirement depends on implementation

Dimension	REQUIREMENT	Ag. Scenario	LBL Scenario	LBL Simu
Precision	HIGH	Depends on asset class	✓	✓
Probabilistic interpretation	HIGH			✓
Tolerance for heterogeneity	HIGH		✓	✓
Granularity	HIGH		✓	✓
Comprehensiveness	HIGH			✓

5.9 Mortgage loan modification program evaluation

Description: Evaluate the impact of several proposed loan modification programs (e.g., those that reduce interest payments, those that reduce principal payments, those that combine these), subject to different eligibility criteria (e.g., minimum coupon, minimum LTV, loan status, etc.). This may be done in the context of a bank's own program or a governmental program.

Key dimensions: This application brings us full circle and back to policy analysis. However, in this case, the key to the analysis is the ability to evaluate the impact of including or excluding certain loans from a proposed program and to then evaluate the impact on the future behavior of those loans that are included. his requires the ability to perform analysis at a granular level and to do so for a wide array of different loan types.

Dimension	REQUIREMENT	Ag. Scenario	LBL Scenario	LBL Simu
Precision	HIGH	Depends on asset class	✓	✓
Tolerance for heterogeneity	HIGH		✓	✓
Granularity	HIGH		✓	✓

6 Discussion

Up to this point in the paper, we have implicitly assumed that users face mutually exclusive decisions; that only a single approach may be chosen for a particular application. However, in our experience, even though there are often clear recommendations with respect to which approach will serve a particular analytic need *best*, when users have the option to do so, we find that a *combined* approach often results in deeper insight than a monolithic one. In this section, we discuss briefly how the various approaches we have been examining can be used in concert to complement each other.

6.1 Combining single-path stress-testing and simulation-based loss distribution analysis

We find it useful for many credit risk applications to conduct both stress testing and simulation analysis. Said more strongly, using one without the other can lead to oversights in risk management.

Scenario analysis provides a concrete, intuitive description of states of the world that might occur and the losses associated with those states under the model. This is valuable for understanding a model and a portfolio and for gaining intuition on the drivers of credit risk for the portfolio. This intuition can be particularly important in understanding a model or a portfolio being modeled.

Simulation-based analysis, on the other hand, is one way to produce a detailed description of the loss distribution for a portfolio. It provides information about the range of possible losses *in a probabilistic setting*. This can be useful in assessing capital requirements and for other portfolio management activities.

Importantly, while simulation approaches provide a wealth of detail, a manager or regulator may be concerned with one or several very specific scenarios. By explicitly running stress tests using such scenarios, users can gain understanding of how the portfolio might behave under such a state of the economy.

Stress-testing can also be used as a reality check on models and portfolios. A user can evaluate losses under a stress scenario and compare these to the simulated distribution of losses. If the losses appear very much higher than under most states of the simulation, the user can use this information to delve into the source of the differences. These might be due limitations of the model or they may arise because of the relative likelihood of the stress case. Similarly, if losses, under what the user considers to be a very stressful scenario, do not appear high compared to those under the simulation, the user can use this information to better understand the drivers of portfolio risk. Using both simulation and scenario-based approaches permits users to combine managerial and analytic judgment with the probabilistic simulation results in a way that is often more satisfying and informative that either one alone.

6.2 Combining aggregate- and loan-level analysis

It can also be valuable to apply aggregate- and loan-level approaches, particularly when policy decisions are involved. As in the case of combining simulation and stress testing, the two approaches can serve different purposes or can serve to provide complimentary insights.

Many retail portfolios, particularly those containing residential mortgages, are characterized by marked heterogeneity and non-linearity. In such settings, a loan-level analysis is useful for understanding how the heterogeneous loans and borrowers lead to layered risks. Differences between aggregate- and micro-level analyses can be marked in such settings and seemingly similar aggregate summary statistics can mask substantial differences in real portfolio risk. However, even when asset structure suggests loan-level analysis, the additional information provided through aggregate analysis can inform decisions that go beyond the individual portfolio.

For policy decisions or marketing strategy involving broad trends in global or regional markets or for assessing how entire markets might change under different states of the economy, aggregate analysis provides a powerful tool. The ability to easily combine aggregate estimates of different national and regional macro-economic forecast to arrive at aggregate forecasts of the behavior of a key asset class can be valuable as a in understanding the ways in which a broad asset class (rather than a specific portfolio) may be changing.

Combining this analysis with an analysis of portfolio concentrations can suggest fruitful areas for diversification, as well as informing business franchise decisions. While many risk managers focus primarily on their own portfolios, they often forego similar analysis of the markets in which they operate and through which they must grow, hedge and trade their portfolios. Combining aggregate- and micro-level analysis can serve to close this informational loop.

Finally, there are settings in which the two approaches may be combined in a single analysis. For projects involving many asset classes or regions, many institutions elect to stress-test disparate portfolios using the same macro-economic assumptions. In cases where loan-level data and models are available for only some asset classes, but where those asset classes are key from a business perspective, institutions may choose to use both aggregate and loan-level analysis in a single exercise, applying loan-level analysis for the key asset classes that and aggregate-analysis for the remainder. The combined analysis can enforce common macro-economic assumptions through the use of the same stress scenarios.

7 Conclusion

We have attempted to provide some clarity on practical model selection challenges and how they may be addressed through different approaches. To do so, we have extended the model selection framework introduced in Dhar and Stein (1997) to accommodate features relevant to retail portfolio analysis. Our goal was to provide readers with the outline of a framework with which to consider their applications and options in business settings.

Table 3, below provides a concise summary of the analysis of our sample applications. We emphasize again that there is nothing about these applications that make them particularly relevant to any given user. They provide a cross section of some of the applications that some users may encounter.

Table 3 Summary of preferred models from examples in Sections 5.1 - 5.9

Sample Application	Typical recommendation
Regional macro-economic stress testing for a single asset class	Aggregate Scenario
Stress testing retail portfolios across bank w/in a common model (loan-level data or models not available/practical for asset classes)	Aggregate Scenario
National economic policy evaluation	Aggregate Scenario
Capital threshold analysis for key asset classes (loan-level data or models not available/practical)	Aggregate Scenario
Capital allocation for key asset classes (loan-level data and models available)	LBL Simulation
Stress testing heterogeneous mortgage portfolio (loan-level data /models available)	LBL Scenario
Investment and loan selection/origination	LBL Simulation
Transfer pricing	LBL Simulation
Loan modification program policy evaluation	LBL Scenario/Simu

In model implementation, there are always trade-offs and these tend to be organization and application specific. We hope that this framework and discussion helps put some structure on these decisions and that the framework provides a starting point for readers to expand or modify in considering their own applications.

8 References

1. Bohn, J. R. and R. M. Stein (2009) *Active Credit Portfolio Management in Practice*, Wiley, Hoboken, NJ.

2. Chinchalkar, S. and R. M. Stein (2010), Comparing loan-level and pool-level mortgage portfolio analysis, Working Paper, Moody's Research Labs, New York, http://www.moodysresearchlabs.com/Research/Default.aspx#WorkingPapers

3. Dhar, V. and R. Stein (1997), *Seven Methods for Transforming Corporate Data into Business Intelligence*, Prentice-Hall.

4. Kumra, R., R. M. Stein and I. Assersohn (2006), Assessing a knowledge-based approach to commercial loan underwriting, *Expert Systems with Applications*, Volume 30, Issue 3, *Intelligent Information Systems for Financial Engineering*, Pages 507-518.

5. Zandi, M. , C. Chen, deRitis , C., Carbacho-Burgos , A., (2009), Housing in Crisis, When Will Metro Markets Recover?, Report, Moodys Economy.com, February, available at: http://www.economy.com/home/products/housing/default.asp

6. Zandi, M. and A. Blinder (2010), How the Great Recession was Brought to an End, Moody's Analytics, West Chester, PA, available at http://www.economy.com/mark-zandi/documents/End-of-Great-Recession.pdf