REVENUE AND LOSS FORECASTING FOR BANKING

Fundamental principles of Corios Forte

by Robin Way

CORIOS REDPAPER — Papers that get read.™

CORIOS **RED**PAPER — Papers that get read.™

REVENUE AND LOSS FORECASTING FOR BANKING

Fundamental principles of Corios Forte

by Robin Way

Scope

This is a primer on the structure and scope of revenue and loss forecasting for financial services lenders and credit issuers, based on Corios' field experience with delivering forecasts for more than ten of these FSIs in the US and Canada.

The scope of this document focuses on the value contribution of improving the process of revenue and loss forecasting; describes the business structure of revenue and loss forecasting; characterizes the account lifecycle dynamics that drive these forecasts, and illustrates implications for the analytic modeling strategy.

This document is not intended to teach a new practitioner in how to become a business forecaster or an econometric forecaster; instead it is intended to provide business content to the broader business team so they are briefed on the issues important to marketing, risk management and finance professionals who produce these forecasts.

Business value of revenue & loss forecasting

Contribution to margin and risk mitigation

There are several strategic ways in which the revenue and loss forecast contributes to the margin and risk mitigation strategies of the bank.

First, a solid forecast contributes to the estimates of earnings that the bank will deliver to investors. Evaluations of the performance of the bank's earnings, and the demonstrated ability of the bank to meet those expectations, have critical impacts on the bank's rating by investment analysts. These ratings in turn affect the bank's cost of funds.

Another essential element of the bank's perception among analysts is the stability of earnings over time; the better that a bank is able to identify and address the sources of volatility in expected earnings, the better their perceived management strength with these analysts. A strong expected growth in earnings that is accompanied by high volatility in past earnings (compared with their forecast) is not as attractive as strong earnings growth that matches expectations.

An internally-focused contribution of the revenue and loss forecast is the contribution to marketing & risk management policy. Adjustments of both the large (strategic) and small (tactical) variety to today's decisions on issues as wide ranging as acquisitions, pricing, exposure, line increases, and collections have significant impact on forward expectations of company performance. Decision-makers overseeing these issues use the forecast as a baseline for making and measuring the impact of the decisions they make on performance in the ensuing months.

Another internal contribution of the revenue and loss forecast is on setting the provision against losses. This can be one of the largest items on the balance sheet that dilutes earnings. Funds set aside to cover expected losses are funds that cannot be used to generate earnings; set aside too large of a loss provision, and earnings may suffer. The alternate case is not attractive either; set aside too small of a loss provision and the bank may incur a reduction in earnings due to not covering their losses. This shortfall might be worse than the alternative of having set aside too much loss provision. In either case, a solid forecast contributes to making prudent decisions on loss provisioning.

Finally, the revenue and loss forecast makes a sizable contribution to the budget cycle, in developing a foundation for making investments in acquisitions, buying portfolios from other issuers and lenders, expanding the bank's footprint, and expansion of new and existing channels for account management and customer relationships.

Initiatives to improve the forecasting process

Most banks create their revenue and loss forecasts starting with basic spreadsheet models, and it is common to find the finance function of the bank operating the revenue and loss forecast process in cooperation with marketing and risk management providing the necessary analytic inputs and data feeds. The revenue and loss forecast might contain components executing as a large spreadsheet model, a database or a SAS application. Typically these kinds of systems aren't designed to be dynamic or flexible to changes in analytic requirements without modifying hard-coded functionality.

There are multiple reasons that a bank would desire to improve this capability. First, there is the need to create better diagnostics within the forecasting system so that forecasts can be improved over time, factoring in the results of variance analysis at the conclusion of each performance period and comparison with the forecast for that period. This leads to the identification and sensitivity assessment of drivers of revenue and loss performance, which often differ in their impact across the layers of the segmentation hierarchy.

The segmentation itself might be a focus for wanting to change the revenue and loss forecasting system. It is common for the segmentation to either be too aggregate or too granular. It is relatively simple to build a spreadsheet-based model that focuses only on a few dimensions of a portfolio-level forecast. However, as more stakeholders invest in the use of the forecast, there is a desire to add more segmentation schemes into the forecast. This can lead to too much granularity as the number of accounts within each cell of the hierarchy becomes too small. This over-granularity can also lead to greater volatility within the observed trends in those market cells.

It is best to find a balance between too many and too few layers in the segmentation hierarchy. Experience suggests this balance should focus on identifying the segmentation design calibrated to those layers that contribute to the largest share of variability in the forecast metrics themselves. For example, it might be common practice to break down the revenue forecast by geographic region, but if all regions share similar growth rates and similar levels of forecast accuracy, choosing the geographic layer in the segmentation scheme won't contribute to the accuracy of the forecast model. Instead, it is more likely that behaviorally-based segmentation layers (such as credit quality or spending volume) will tend to contribute towards accuracy more substantially.

A benefit of designing an analytical accuracy-based segmentation scheme is to reduce the uncertainty in the forecast. All professional forecasters are familiar with the conventional (if initially disarming) belief that "all forecasts are wrong". This belief is interpreted to mean, any point forecast that a forecaster provides to a decision maker will almost never be right (i.e., to the specific dollars and cents), but the value of a good forecast is in providing a range in which the actual performance metric will likely result, preferably with a statistical measure of certainty (such as a confidence limit). The smaller that this confidence limit is (assuming the forecaster has done his or her job correctly, and assuming the past is a good predictor of the future), the lower the implied volatility in the forecast, and hence the reduced operational risk to the bank of making a decision on a forecast with broad volatility.

Finally, decision makers typically use the forecast as a baseline against which to compare alternative ways of improving future performance by introducing new portfolio management strategies or changing existing strategies. The more stable the forecast, then the decision makers will have a more robust belief in the baseline against which to measure the contribution of these management strategies.

Stages of forecasting process maturity

As a bank adopts a new forecasting approach, they will evolve through various stages of learning and maturity. Based on previous experience, these stages might resemble the following description.

Business as usual (“BAU”): Client has traditional forecasting tools (Excel, SAS code), and wants to improve accuracy, quality of diagnostics, actionable outcomes of the forecast results. Internal stakeholders want more flexibility to changing information needs, more relevant forecast breakdowns, better ability to understand why the forecast overshot/undershot the actual performance so they can provide a clear explanation to partners and third parties.

Stage 1: try to expand the use of traditional tools to meet these expanding information needs. Result: spreadsheets are fragile; all the formulas needed to maintain the forecast system get corrupted as the dimensions of the forecast change; custom coded applications become more and more hard-wired and incomprehensible, and more dependent on the small team of experts who can keep it running. More time is focused on the functions and less on the accuracy and analysis.

Stage 2: The forecasting team’s management decides to evaluate and introduce new forecasting system tools. Using this as an opportunity to overhaul the process and meet all requirements, the team and partner identify the ideal segmentation, forecast drivers, forecast metrics, sensitivity analysis, and analysis & interpretation needs. Start rebuilding the forecast system to meet these new needs and produce initial results of the forecast. Monitor BAU forecast results and identify variances, new insights.

Stage 3: Management receives initial forecasts and within a few months reviews the forecast results in comparison to actual performance. Management starts to consider changing their decisions and policies in response to improved information.

Stage 4: Forecasting team, following several months of forecast-actual comparisons and variance analysis, has enough information to monitor improvement in business metric performance, such as loss provisions, earnings stability, forward expectations, further calibration of forecast system and assumptions. The forecasting team implements adjustments to forecasting assumptions and starts executing proactive what-if analysis and stress tests on the forecast.

Stage 5: Management starts believing in the new forecasting results, and asks for the forecasting team to start incorporating judgmental adjustments to the forecast that represent forward-looking strategic adjustments to business plan, such as above-the-line adjustments to forecast given risk, acquisitions and portfolio marketing policy initiatives. New forecasting system takes root in the operations of the bank.

Revenue & loss forecasting best practices

This is an important section of this document. This content is what makes the difference between a generic statistical forecasting tool and a business-relevant revenue and loss forecasting engine for credit issuers and mortgage lenders.

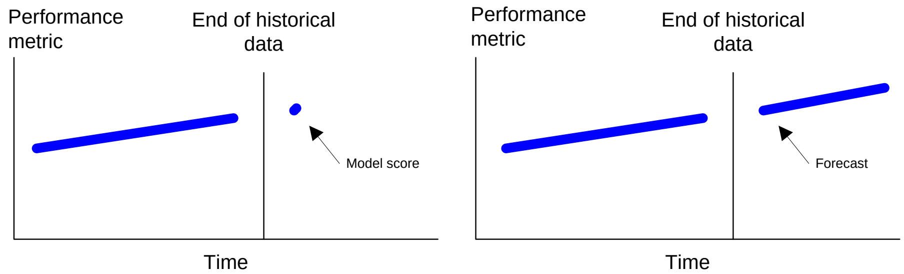
Segmentation granularity

The proper segmentation of accounts provides the opportunity to improve the accuracy of the forecast and to produce actionable insights for decision makers. Designing and executing the segmentation properly is essential to success of the forecasting process.

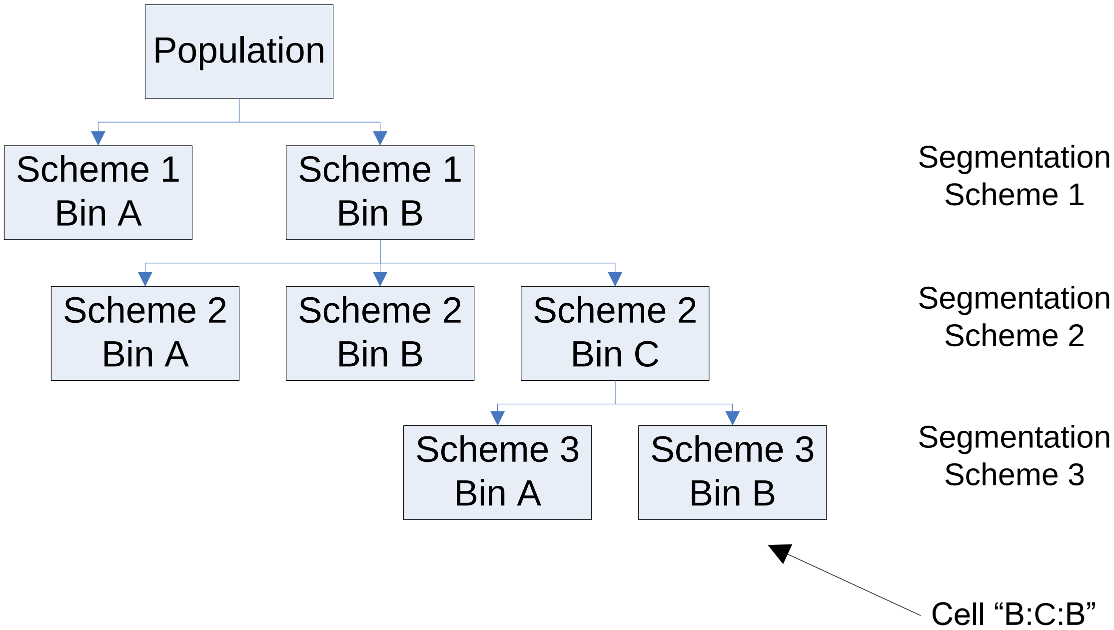
First, why use segments at all? Why not just forecast at the account level? After all, it is a common practice to build statistically-derived model scores at the account level for acquisition, booking, response, delinquency and collections at the account level. Why not use these scores? Well, these model scores are based on a snapshot of historical account-level performance and a prediction of that behavior at a specific point in time, and not for the trend taken across discrete points over a period of time.

More specifically, the outcome to be predicted by a scoring model is expected to occur within a specified interval, such as “in the next 3 months” or “in response to this type of offer”, and not to build a statistically reliable projection of account performance for multiple periods for an extended interval, such as “each month for the next 12 months”. In order to build a reliable projection over time, we aggregate accounts into groups, and project performance either for the entire group of accounts, or we distill the performance of the “average” account for that group and apply the projected average performance to all accounts in the group.

Comparison between a model score and a forecast



The segmentation used in a revenue and loss forecast is typically represented hierarchically, consisting of multiple levels, with each level containing a single segmentation scheme. Each scheme consists of bins. The top of the hierarchy represents the population of all accounts across all segments. At the bottom of the hierarchy, each set of accounts in a unique combination of bins is called a cell, named by the unique values of each bin contributing to its definition.



Commonly-encountered segmentation schemes include

- credit quality,
- business partner portfolio (e.g., a specific airline, a specific retailer),
- card product type,
- purchase APR bin,
- wallet share,
- transactor/revolver behavior,
- frequency / recency / monetary spend behavior,
- balance / payment behavior, and
- dominant merchant category spend.

As you can see there are quite a few potential segmentation schemes that could be of interest, and each scheme could consist of anywhere from two to 20 bins. The permutations among these schemes and bins are potentially vast.

The forecaster needs to strike a balance between diagnosticity (i.e., using a large number of schemes and bins in the segmentation) and stability (i.e., having enough accounts in each cell to permit the signal to be stronger than the noise). A best practice in segmentation design with regards to revenue and loss forecasting is to identify those segmentation schemes and bins that contribute to variance in key metrics such as average monthly balance or delinquency ratio. Sometimes a segmentation scheme will be strongly related to variance in these metrics but not all bins within that scheme have equal contribution to that variance; in cases like this, some of those bins can be combined together and treated as a single bin. This preserves stability without sacrificing diagnosticity. Statistical approaches like analysis of variance and data mining approaches including information reduction assist in the identification of the key schemes and bins that are most analytically desirable.

Not all metrics need to be tracked at the most granular level of cells. Based on the analysis in the segmentation design, some metrics like the number of active accounts may be stable down to the lowest cell level, but other metrics might be measured and forecasted at higher levels of aggregation, such as finance charges per active account.

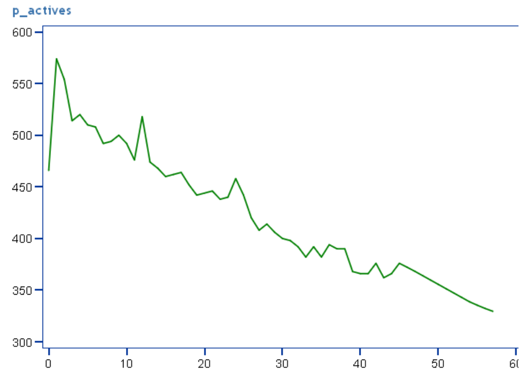
The segmentation should strive to maintain a minimum number of accounts per cell. There are both simple and sophisticated ways of determining what this minimum cell size should be. Simple ways, such as a rule of thumb that no cell should contain less than 500 accounts, can be perfectly satisfactory; it depends in part on the underlying stability of the trend and the desired precision of the forecast estimates. Statistical power calculations, such as those calculated by SAS' Power procedure, can be used to refine those minimum cell size estimates, perhaps on a per-metric basis.

Account lifecycle

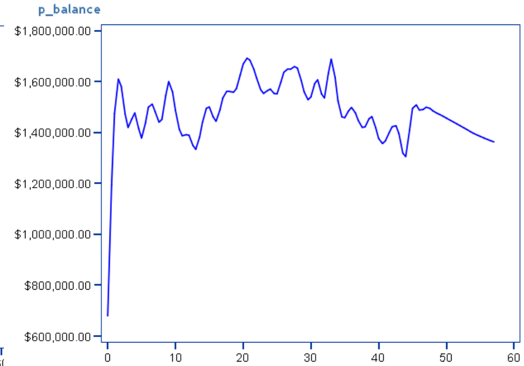
There are some generally-accepted stages of the card account lifecycle: application, booking, activation, balance growth, maintenance, delinquency, paydown and closure. The transition points from one stage to the next can often be tied to specific ranges of the time that passes after the account is booked; when expressed in months, we call this account lifecycle measurement "months on book" ("MOB").

Some metrics, when you look at them over time, reflect some underlying trends which tend to look similar on the average across accounts in a segmentation cell. Accounts when booked start to build balances in MOB 1-3, followed by a gradual slowdown in balance growth over MOB 4-6. This balance growth will start to decline gradually over the next 20 to 30 months. The number of active accounts will begin at zero in MOB 0 (the month of booking) and then jump to their maximum level in MOB 1-2. The rate of active account attrition starts right away and the reduction in active accounts is typically a fairly stable rate of decline. Account delinquency starts to ramp up in the months following balance build, and the trend of accounts starting to occupy the higher DPD bins follows a structured rate as MOB's continue to grow. Charge-offs begin to accrue in MOB 7 (account activation plus 180 days) and grow through MOB 30, by which point the standard wisdom holds that most accounts that will charge-off have done so already.

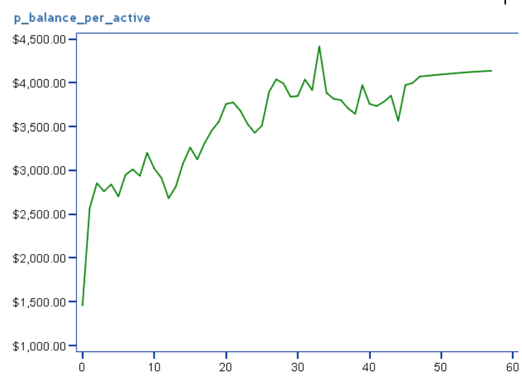
Active Accounts



Balance ($)

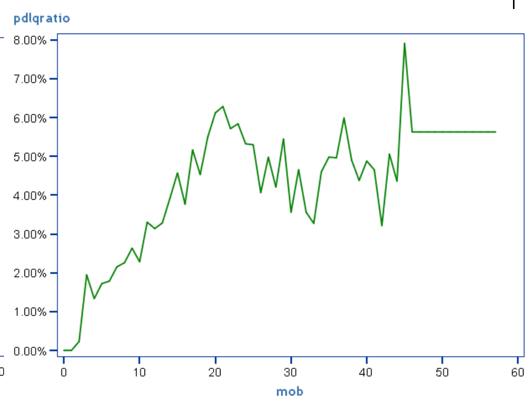


Balance per active account ($)

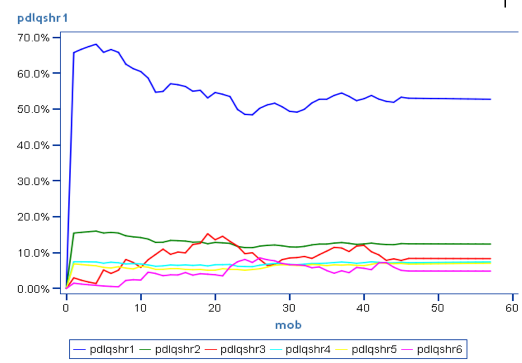


Delinquency ratio (%)

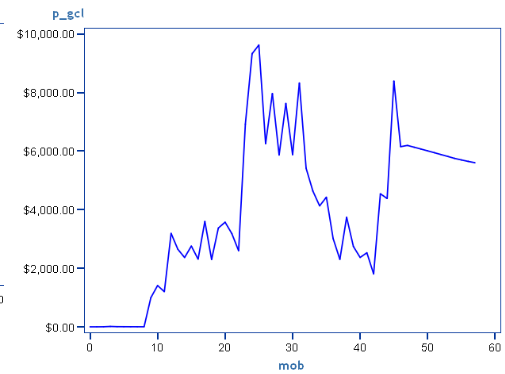
(all delinquent accts / all active accts)



Delinquency rates (%) by DPD bin (1 – 6)



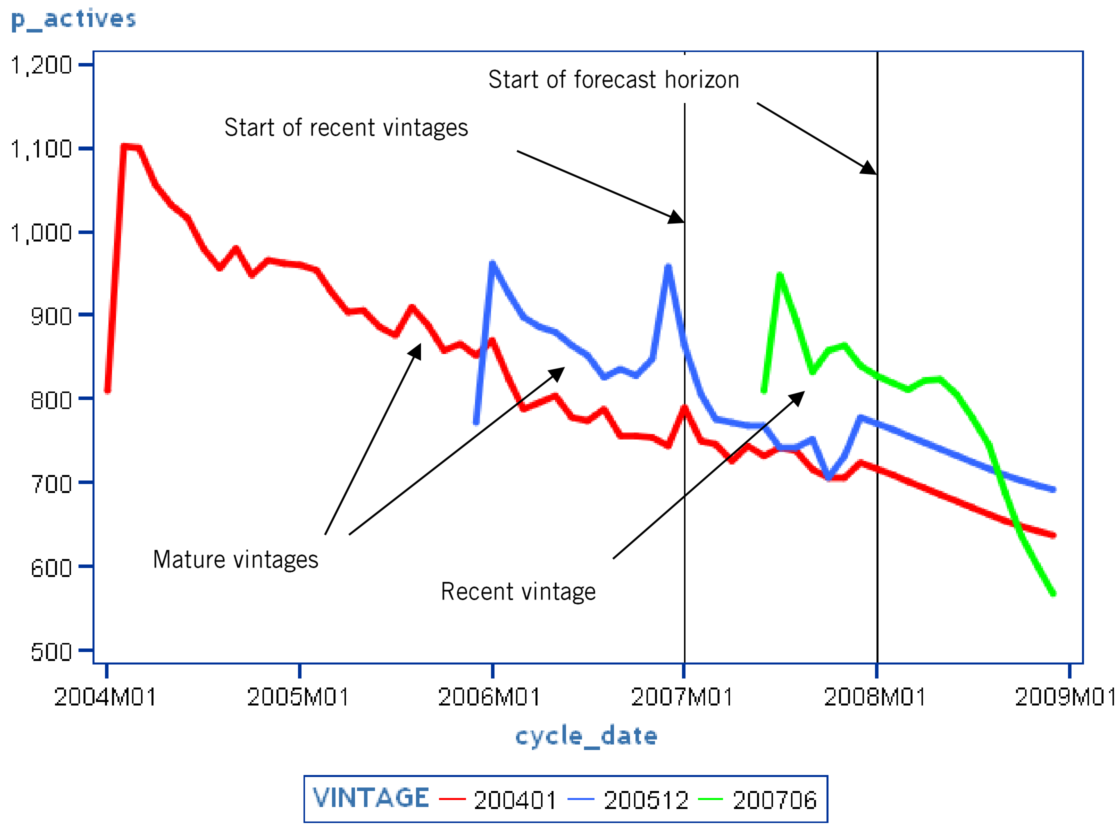
Gross credit losses ($)



Some metrics are directly under the control of the issuer, such as application, booking and involuntary closure. Other metrics are mostly outside the issuer's control, other than influencing them through marketing and risk strategies. These include activation, spend, balance growth and paydown. If the metric is under the direct control of the issuer, it might be more appropriate to use assumptions from past performance and build multiple forecast scenarios using alternative assumptions on these metrics.

Vintage maturity

Directly related to the concept of account lifecycle is the concept of vintage maturity. Each account is a member of a vintage (see above in the "Business Dimensions: Time" section for the definition of "vintage"). As the vintage ages, we learn more about the trends inherent in the vintage's performance across each metric of interest. While there is no hard and fast rule for the number of months on book that constitutes vintage maturity, it might be acceptable to consider a vintage that is 12, 18 or 24 months old as relatively mature in terms of having established a stable trend on most metrics.



However, there is a conundrum in producing a revenue and loss forecast with regards to vintages: with each passing month, the bank will be taking on more and more new vintages, and the performance of these vintages has to be taken into account for the forecast to be accurate. But we have no information on these new vintages, because at the point of producing the forecast, we know virtually nothing about the accounts that make up these vintages (because we haven't acquired them yet). For example, if it is now October 2008, how much do we know about the accounts we will bring on book in June 2009? (Answer: virtually nothing.)

Moreover, vintages that today are not yet mature also have to be forecasted, although their own performance trends have not yet stabilized. Using the same example from the paragraph above, how should we forecast performance for the accounts in the August 2008 vintage? We know how many accounts were booked, but these accounts have only started to activate and build some initial balances. We'll have basically no firm information on likely delinquency rates for this vintage until well into next year. How can we tackle this thorny problem?

One approach is to take what we know from the performance of already-established mature vintages and apply these performance trends as a projection of how recent and new vintages will perform. We make the assumption that the mature vintages we use are the best indicators of future performance of new and recent vintages. To do this well, we should identify the factors that the mature vintages share with recent and new vintages—such as being members of the same segmentation cell(s). If the basic attributes of a segment have been fairly consistent over time for past vintages, perhaps those trends will continue and we can use those assumptions for new and recent vintages.

The key to applying trends from mature vintages to the new and recent vintages is to tie them to the month on book as an index. This relates directly back to the notion of the account lifecycle. This is also why we calculate derived metrics like balance per account—the balance of any historical vintage in total dollars will be very different from the balance of a new vintage, but the balance per account might be a decent proxy that we can apply to the new vintage, as long as we know or can project how many active accounts will exist for that new vintage from its inception through its maturity. This can be done in a straightforward way for recent vintages because we already know how many accounts we have and we can forecast the trend of active accounts for those vintages. For new vintages, we can consult with the acquisitions planning team to determine the most likely number of acquisition contacts, and then use some historical estimates to determine likely applications, bookings and activations.

There are a few nuances to this analysis to bear in mind. Credit issuers often acquire portfolios from other issuers that already have pre-matured vintages on book. When this data shows up in the historical performance file, you may find that the month on book field for these mature vintages does not start at zero, but instead some number larger than zero. This could be because the new issuer didn't warehouse the historical performance for these vintages in their own analytics dataset. In these cases, you need to exercise caution in using this raw data in a statistical forecast; instead look for proxy vintages from other segments, possibly from non-acquired vintages.

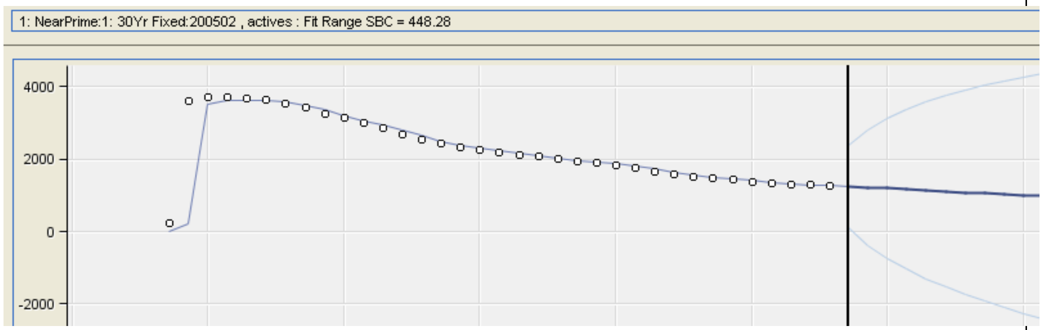
Another nuance you may run across in historical data is what one could call "super-mature vintages". If the issuer has been in business for some time, there may be some vintages that are over 10 years in age or more. In this case, one approach is to consolidate the long histories for these vintages into a single aggregated cell and treat it separate from the rest of the month-on-book forecasting approaches. There are a variety of ways to handle this that go beyond the scope of this paper; the key is to recognize the possibility that these kinds of super-mature vintages exist.

Estimating, storing and assigning curves

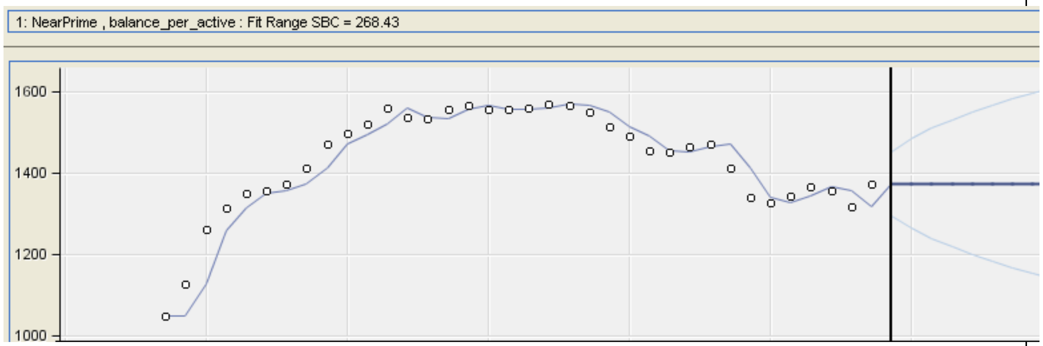
As described in the section above, we use mature vintages to estimate statistical models that are intended to fit the historical performance and project the future performance of each metric. The results of these statistical models are captured as curves that represent the (optionally smoothed) percent change month-over-month, indexed to the month on book. The use of these curves on recent and new vintages allows us to forecast the future values of all relevant metrics for the revenue and loss forecast.

While there are multiple ways of estimating these curves, the statistical forecast is a reliable approach that can be performed (through the use of SAS Forecast Server) at scale for hundreds or thousands of cells in the segmentation hierarchy.

Statistical forecast of Active accounts



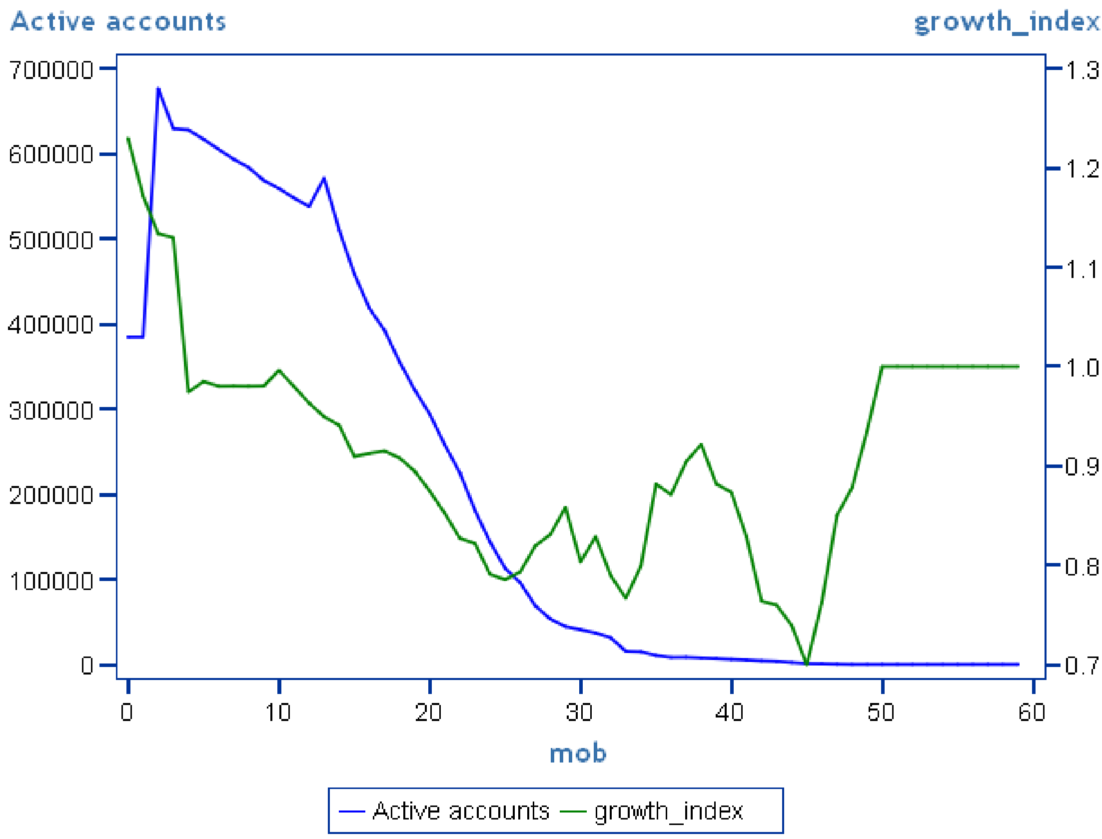
Statistical forecast of balance per account



There are other ways of estimating curves, such as using time-dependent curve generation routines, but these generate generic curves that require a second pass of manual tuning of the curve to fit the actual historical data. Since the selection of the initial pass (to generate the selected curve and its parameters) and the second pass (to tune the curve to the data) are both manual processes and subject to considerable judgment by the forecasting analyst, one could conclude this approach is not scalable to a large number of segmentation cells.

Let's return to our discussion of using statistical forecasts to generate the projections used to produce curves for each metric. These metric trends by month on book can be represented as a "curve", both in terms of a visual curve on a plot (like the ones in the figure above) and as a conceptual representation of the growth rate of a metric over time. When we create these curves, one of the best ways to physically calculate it is as a percentage change from the previous period. So if we know the initial value of the number of active accounts in month on book 1, and we have a curve that represents the percentage change in each period for the number of active accounts (from past mature vintages), we can apply those calculations forward and develop a projection of the number of active accounts in every month on book going forward.

We use mature vintages to construct the model-based raw curves by segment and vintage in both the historical and future period. These are raw curves because we haven't yet converted them into percentage change yet. We might also want to first smooth these raw curves out a little (i.e., using a moving average, via SAS' Expand procedure) before calculating those percentage changes, to avoid wildly fluctuating percentage change indicators, especially when the metric is small in number and the percentage change calculation is sensitive to relatively large changes in the denominator term. The “percentage change” series becomes our final curve for application to the recent and new vintages.



In the revenue and loss forecasting engine that is used in the RBD demonstration, we store all these curves for revenue metrics in a single data file (and loss curves are stored in a second file). The curve for each series at each level of the segmentation hierarchy for mature vintages is stored in this file, tagged with a curve descriptor that contains the concatenated values of the segmentation level and the vintage, by metric and month on book. Note that we are also storing top-down and intermediate-level cell-specific curves, because we may deem that some of the vintage-specific or bottom-level cell-specific forecasts are not reliable enough to use broadly.

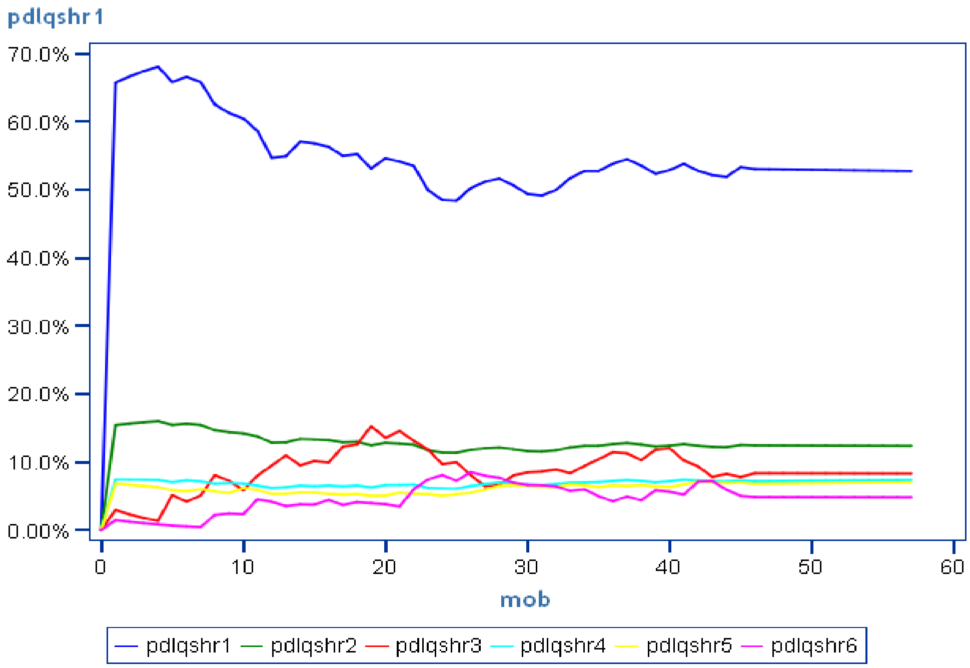
Once we have stored these curves, we can assign any curve for any metric to the projections for recent and new vintages. A series of business rules can be used to perform this assignment, or manual decisions can be made to specific segments and vintages among the new and recent vintages. This activity is part of the functionality in the forecasting engine.

Note that because we store these curves by month on book, this can tend to ignore seasonality that is calendar-specific, such as spend spikes in the holiday season. If the metric is highly calendar-seasonal, you may want to note that, and assign a curve from a historical vintage with the same start month in a previous year. This will help align the spending spike with the appropriate calendar month for the new or recent vintage.

Delinquency modeling

Modeling the forecasted delinquency and charge-off performance for a bank is a prime area of interest. There are a variety of approaches available to model this series of metrics, some of which are straightforward and others quite sophisticated. We can promote an approach which is novel but has undergone some initial evaluation in the field. Its primary attractiveness is that it can accommodate forecast drivers and it is scalable as it draws on a statistical forecasting technique native to SAS Forecast Server.

Recall that delinquency is a situation where an account can reside in a variety of delinquent states, indicated by DPD bins 1 through 7 (where 7 represents charge-off, a state from which they cannot return).



One simple approach is to take a snapshot of historical data and determine the per-period transition likelihood from one bin to another. Envision a grid where the first axis lists the bins in which accounts resided last month, and the second axis lists the bins in which accounts resided this month. The numbers in each cell of the matrix include the count of accounts and the percent of accounts of the row, column and complete table. This is sometimes called the "is-was" matrix. Some issuers simply use the percentages in each cell at that snapshot in time to estimate the transition likelihood for migration from the previous state to the current state. An expansion on this technique uses the same statistics but models the transition likelihoods using a technique called Markov Chains. The advantage of the Markov approach is that it can be used to project changes in the transition states over time, lending itself to more of a true forecasting approach.

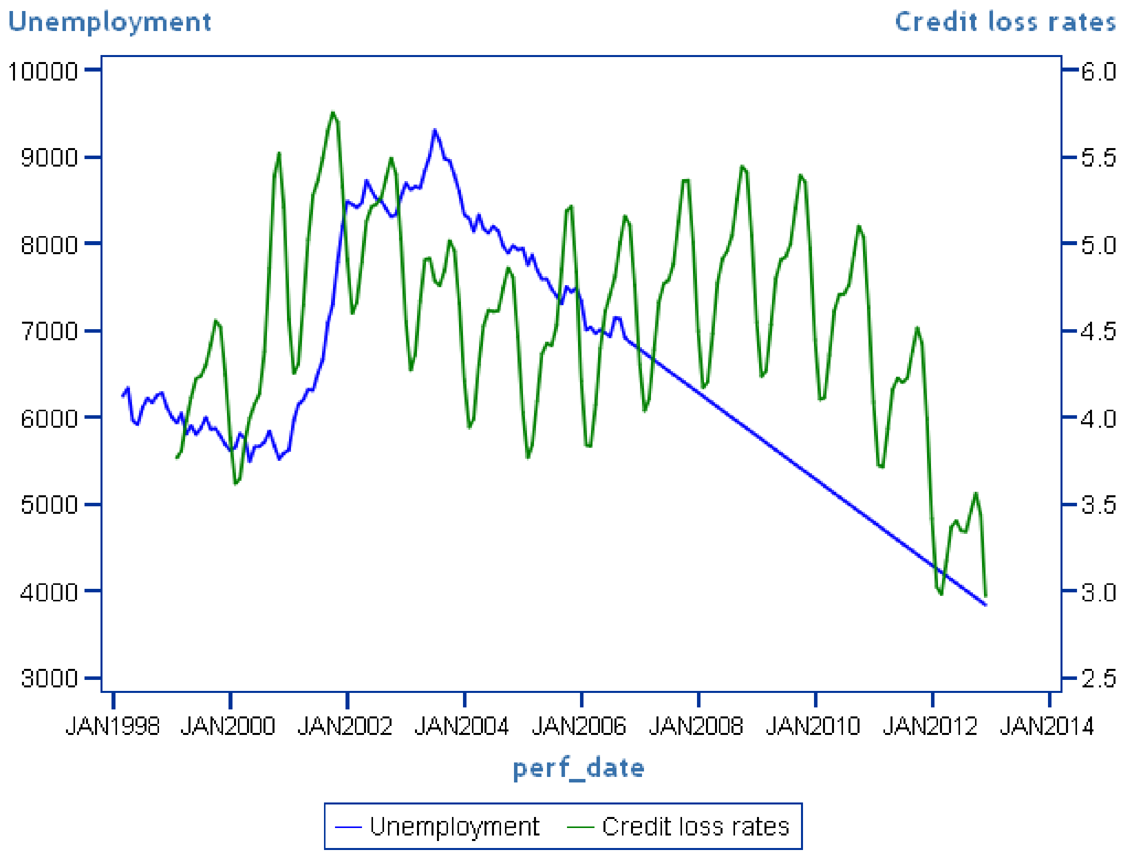
Yet another approach that uses a more sophisticated econometric technique, and can incorporate drivers such as macroeconomic indicators, is called generalized logit regression. Instead of forecasting the transition likelihood of every cell in the "is-was" matrix, it just looks at the distribution of accounts in each DPD bin at each month in the historical period. The percent of accounts in DPD3, for instance, if mapped out over time, form a trend line that could be forecasted. By using economic drivers to augment this model, changes in expected economic conditions can be used to influence the forecast. An advantage of this approach is that all DPD bins can be forecasted simultaneously, but a downside of the approach is that these models are fairly sensitive to changes in historical trend or changes in the value of the economic drivers. This means that they need a lot of attention by the forecast analyst on an ongoing basis.

The approach supported by Forecast Server shares some similarity with the generalized logit approach in terms of the structure of the historical data, and the ability to use economic drivers. However, in this case, each DPD bin's likelihood is estimated as a separate forecast function. The use of the Forecast Server reconciliation function is used to force all likelihood scores across all DPD bins to sum to 100% in each period for each series.

While this section focuses on delinquency likelihoods, a related topic that can use the same approach is segment migration. Segment migration refers to the change in behavior over time by a set of accounts, for example from transactor to revolver status, or from low risk to medium risk in terms of payment behavior. Unlike with delinquency, there is not usually a sequential structure of migration from one state to a neighboring state when considering behavior migration; however the same technique can still be used to tackle this analytic problem.

Forecast drivers

A forecast driver is a factor that influences the historical and future performance of the bank's portfolio. Some of the most common factors in a revenue and loss forecast are macroeconomic drivers, such as bankruptcy rates, inflation rates, federal funds rate, prime lending rate, national credit and mortgage delinquency and charge-off rates, national credit availability, national credit quality, national income growth, and cost of living indices.



A bank might be interested in using factors that reflect its own strategic decisions, such as changes in acquisition spending or credit granting policies or collections policies. However, these factors are different than macroeconomic series because the latter are outside of the bank's control and there is some uncertainty about the future values of these series. Hence the way the economic series gets used in the revenue and loss forecast is more likely to be rooted in a statistical forecasting approach. The use of bank strategies and changes in policy are more likely to be modeled as a scenario, applied as a judgmental override to the forecast of one or more metrics. Forecast Server supports the judgmental override function directly.

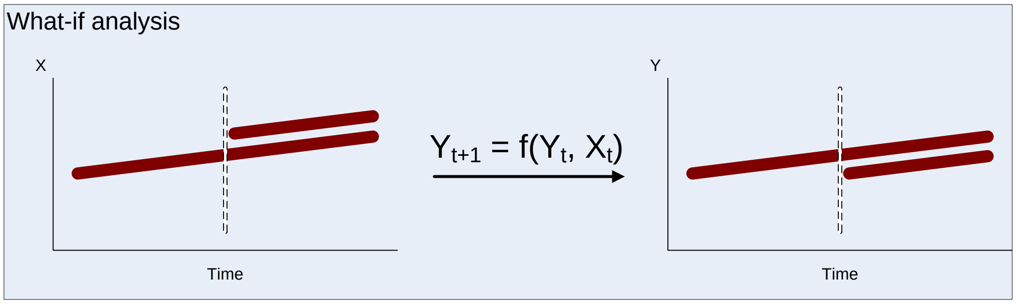
Consider the role played by unemployment in a forecast of revenue and loss. In a situation where unemployment rises, the ability of individuals to pay their bills, using all forms of payment, becomes impacted. Some individuals will rely on their revolving credit to get past a situation of brief unemployment, while others will cut their discretionary spending as much as possible. Accounts that regularly pay more than their minimum monthly payment amount may change their payment behavior. It is likely that individuals will choose to pay only those bills that keep them fed and in their homes, while choosing to become delinquent on other forms of credit.

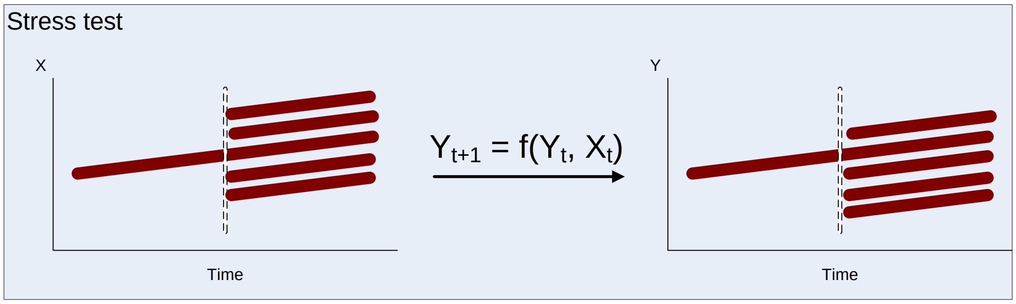
The key to incorporating these forecast drivers in the revenue and loss forecast is to focus on the expectations for the relative value and trend in each relevant economic series in the forecast horizon. While the historical values of these series are important to understand, what is ultimately more important is to understand the future expected value of the series, whether the trend for the economic series will be up, down or stable, how steep this change in the trend (if it arrives) could potentially be, and how sensitive the forecast for any specific metric will be in relation to changes in this and other economic series.

A first step in the revenue and loss forecast in each forecast publish period is to update the forecast of the economic series. This is often a service provided by a third party provider such as Moody's, although some issuers will generate their own forecasts of some of these economic series for their own purposes as well. Once this economic history and forecast file is available, it will be appended to the historical performance data by performance month.

In the context of the statistical forecast (again, we're talking about mature vintages), the macroeconomic series play the role of influencers on the forecast of the metric of interest. If we are predicting spend, and the interest rate is expected to trend downwards in the next 6 months, that might be used in the forecast to indicate a potential increase in credit-based spending, as the cost of using credit will be expected to fall. The degree to which the forecasted spend will increase is a function of the trend in spending by itself, combined with the correlation over time of spending with historical interest rates.

Because the bank cannot control the future trend of these economic series, it is of great interest to most forecasters to pose questions like, "…if interest rates rise 3 percent per month faster than expectations for the next 6 months, what will be the resulting change in spend?" This is a classic instance of what is called a "what-if analysis" in the context of a revenue and loss forecast. Nobody knows exactly what the interest rate trend will be six months from now, but you can come up with plausible situation that reflects a possible future state of the economic indicator(s) and determine via the statistical forecast model the likely result from this alternative scenario, relative to the baseline forecast. A similar analytic exercise is called the stress test; simply put, a stress test is like a what-if analysis, except that it is repeated hundreds of times, with small changes made to the expected trend in the economic indicator(s) and the resulting range in the expected value of the metric(s) of interest.





Credit card business specifics

For a credit card issuer, the dimensions and metrics involved in revenue and loss forecasting are fairly standard. The RBD revenue and loss forecast features most of the commonly-encountered metrics; metrics that do not appear in the demo are due to not being supplied by the client who provided the sample data. The key elements of the forecast are facilitated well by the metrics available in the demo, and if new ones are required, they can potentially be simulated by the presales consultant on a one-off basis.

This section discusses the dimensions, core metrics and derived metrics common to a revenue and loss forecasting system, many of which are included in the RBD demo. This section also discusses the conventions of interest to the forecasting analyst and business management using the forecast.

Business dimensions

There are typically four sets of dimensions in the revenue and loss forecast: account segmentation, time, forecast publication date, and forecast scenario.

ACCOUNT SEGMENTATION

This is the most analytically rich dimensional set. Any revenue and loss forecast can contain several dimensions, and these dimensions can be represented as one or more logical hierarchies. This section discusses the dimensions used in the RBD revenue and loss forecast demo; other common account dimensions are described in the section "Advanced Topics" below.

The "portfolio" dimension is often used to classify the class of credit quality, such as super-prime / prime / near-prime / sub-prime. Accounts with a prime class of credit quality will have different delinquency and charge-off behavior (i.e., both the average level and the volatility over time) compared to sub-prime accounts.

The "segment" dimension is often used to classify the partner whose branded card is held by a set of accounts, such as a specific airline miles card, a specific loyalty card, a specific retailer private label card. Account holders of an airline's elite loyalty club will have different spend and delinquency behavior compared to account holders of a private label brand for a mass market retailer specializing in discount apparel.

TIME

The time set of dimensions has more members than one might expect. Time dimensions typically include performance months, cycle months, fiscal months, operating days, vintages and months on book.

Performance months are typically the same as calendar months (e.g., January 2008), and assume the month starts on the first calendar day of the month. Most revenue and loss forecasts are aligned with performance months because they are coordinated with financial statements published by the bank.

Cycle months are similar to performance months, but track account holders on different statement cycles. Typically card holders are separated into processing groups whose cycle month-end balances, finance charges and late fees are assessed, and statements generated, jointly on a given business day of the month, or on a revolving cycle throughout the year (such as, 13 statements generated every 28 days). While cycle dates are important to the issuer from an operational perspective, the behavior of card holders with respect to spend, balances and payments tends to not be driven by the issuer's cycle time, but instead by the time of the month the cardholder gets paid and by routine cycles such as filling up the gas tank, buying a large batch of groceries, or longer seasonal cycles such as the winter holiday gift-buying season (followed by the balance pay-down season in January and February).

Fiscal months condition the way forecast results are published, and also tend to drive some management decisions around acquisitions, marketing programs, risk strategies such as collections intensity, and the establishment of annual operating plans and fiscal goals. Hence the nature of the forecast, and especially adjustments to the inputs and outputs, can be heavily conditioned by the fiscal time dimension.

Operating days per performance month could as easily be considered a metric as a dimension, but it plays an important role as an adjustment to other metrics that varies with the performance month. The number of operating days for the performance month impacts the calculation of finance charges and late fees as well as the number of statement cycles that can be completed, and the number of days on which acquisitions, bookings and collections can be processed by the bank.

The ***vintage*** dimension plays a vital role in the revenue and loss forecast because it (along with months on book) captures the account lifecycle dynamic. Banks use the concept of vintage like demographers use the concept of a cohort: a vintage is a set of account holders that were acquired and booked in the same year and month, similar to how a cohort is a set of individuals that were born in the same year and month. For instance, every account that was booked in January 2008 is a member of the "January 2008" vintage and will remain that way until they are retired. Most credit card accounts within a vintage follow a relatively predictable pattern over their lifecycle, broken into conceptual stages like booking, balance build, balance stability, balance paydown, and retirement. The subset of delinquent accounts also exhibit similar patterns that are a function of account lifecycle, such as patterns of delinquency and charge-off. Over time, banks make strategic changes in their acquisitions, booking, account management and collections strategy, and this leads to changes across vintages or more likely, families of adjoining vintages. It is important to stratify the revenue and loss forecasting process to account for the unique qualities of vintages and groups of vintages.

The ***month on book*** dimension is also tied to the account lifecycle dynamic: it is the link between vintage and performance month. Using the example from the previous paragraph, in performance month January 2008, the members of the January 2008 vintage (which were just booked) are in month on book zero (aka "MOB 0"); in performance month February 2008, the members of the January 2008 vintage are in MOB 1. We tend to consider the month in which the account was booked as MOB 0 because the account holders have, on average, only 15 performance days to receive their card, read their terms and conditions for use of the card, activate it, put it in their wallet and start spending with the card. It is more likely to see balance build start to ramp up in MOB 1 through 6. Similarly, delinquencies tend to start ramping up only after MOB 2 (because few accounts have revolving balances yet that they've failed to pay back their minimum amount), and charge-offs tend to start ramping up after MOB 6. Most card issuers forecast performance of all accounts to at least 36, 48 or sometimes 60 months on book, with the assumption that after a certain period of time on book, most accounts that will become seriously delinquent and charge off will have already done so.

PUBLICATION

Forecast publication date is a special instance of a time dimension that merits its own category, because it is used in a different way than month on book or performance month. The publication date dimension is used to capture a snapshot of the historical and projected future performance for the bank at a given point in calendar (performance) time.

The publication date dimension is used to represent the ongoing maintenance and refresh process of a forecasting system over time by the card issuer. Each month, on a fairly stable cycle, the bank will perform the following tasks:

- collect refreshed input data representing the latest month's actual historical performance,
- compare the forecast from the previous period with the most recently-collected actual results (called variance analysis),
- note what analytical adjustments need to be made to the next period's forecast assumptions,
- refresh the scenarios for the next period's forecast,
- perform quality assurance on the results,
- select the champion scenario (and possibly the top challenger scenarios) to present to management
- present initial forecast results to senior management and stakeholders, and gather their recommended adjustments,
- incorporate management adjustments into the champion baseline forecast, producing the final forecast, and
- publish the management-adjusted final forecast.

From the vantage point of the revenue and loss forecasting team, the stages of this process need to be carried out routinely every month, without fail, and with as high a level of quality as possible. The variance analysis step in particular is a focus for the analytical members of the team, as they use the results of the variance analysis to make decisions to tune and optimize the forecast.

Another way to think about the publication date dimension is that it focuses on the judgment and decisions of the bank, as opposed to "external" time such as calendar time or account lifecycle time that is largely outside the direct control of the bank.

SCENARIO

The scenario dimension is used to identify and associate common sets of assumptions on the forecast inputs and intermediate results with a single instance of output forecast results. Banks will generate multiple scenarios in order to evaluate the effect of future changes in a variety of factors (e.g., bank marketing and risk strategy, competing bank strategies, shifts in the economy, and structural changes in consumer behavior and credit quality) on projected revenues and losses. By tracking multiple scenarios over time, the forecast analysis team can compare the sensitivity of historical performance to changes in these factors and use these estimates of sensitivity to gauge likely changes in future performance.

Core metrics

ACCOUNTS

There are multiple classes of accounts, but most credit issuers focus on active accounts when generating revenue and loss forecast results. Active accounts (the most likely metric to be used in the revenue and loss forecast) are a subset of open accounts, and they have activated their card (usually over the phone after being booked by the issuer and receiving the card in the mail) and used the card for a sale at least once in the last M months (where the number M is set by the bank's policy; a likely number is the last 6 months). There are multiple types of active accounts, such as purchase active (used the card for a sale) or debit active (used the card for a sale or a balance transfer or cash advance) and other classifications as well.

Inactive accounts and closed accounts are important to track for the purpose of revenue forecasting since active accounts eventually become inactive in terms of generating new income for the bank, and many of these eventually become closed accounts. Inactive accounts usually get tracked by a behavioral segmentation approach, and the set of inactive accounts could be expected to become reactivated through a marketing offer. Closed accounts can become closed voluntarily by the cardholder, or by the issuer for various reasons, such as entering into collections.

There are other account-like metrics, such as the number of cards in force (since each cardholder can actually have more than one card by an issuer, for multiple members of the household), but usually all income gets rolled up to the account level for forecasting purposes.

SALES VOLUME

A sale results from using the card as the form of payment for goods and services at the location of a merchant who accepts the card, followed by the transaction being authorized, cleared and settled. The "volume" of the sale refers to the dollar amount of that transaction.

OUTSTANDING BALANCES

Cardholders using revolving credit accounts (as opposed to charge accounts like the traditional "green" card offered by American Express) can choose to hold a part of their unpaid balance on the account, which is called the outstanding balance. At the end of the cycle, these outstanding balances generate finance charges, which is a source of income for the issuer.

FINANCE CHARGES

There are at least three types of finance charges from which an issuer can recognize income. Some accounts pay off their balances every month, and generate no finance charge income for the issuer (these accounts are labeled "transactors") and other accounts regularly accrue finance charges (called "revolvers").

The most common finance charge is for outstanding balances generated by sales volume that is revolved at the end of the account's cycle month. The finance charge amount for these transactions is called the Purchase APR.

Issuers also make balance transfer ("BT") offers available to cardholders, which gives the cardholder the ability to move an outstanding balance from another issuer's account (called the "off-us" party) to the issuer making the balancer transfer offer (called the "on-us" party). Issuers make these BT offers attractive by attaching smaller finance charges (called the BT APR) on these transferred balances for a period of time, typically six months. BT finance charges are another way in which the issuer recognizes income.

Credit issuers also allow their cardholders to take cash advances against the card account (i.e., treating the credit card like it's a debit card, with an important exception). These cash advances incur their own finance charges, called Cash Advance APR.

An interesting sidebar is that the issuer will typically allocate the payments made by an account against the APR amount with the lowest APR first. This reflects the importance of finance charge as a source of income for the issuer.

LATE FEES

The cardholder is expected to pay the minimum payment amount each month by the end of the cycle. If this minimum payment is not received, the issuer can (and usually does) generate a late fee on the account. This late fee can either be a flat amount or it can change with the size of the outstanding balance.

INTERCHANGE

Interchange is a fee paid by the merchant accepting the card as the form of payment on every transaction they participate in. The interchange fee consists of a small fixed fee plus a variable fee tied to the sales volume of the transaction. Typically the interchange fee (called the discount rate) is 39 cents for a 100 dollar transaction, but this varies depending on the relationship between the merchant and the payments processor (e.g., Visa, Mastercard, American Express, Discover). This interchange fee gets split between the issuing bank (who "owns" the cardholder relationship), the acquiring bank (who "owns" the merchant relationship) and the payments processor (who "owns" the responsibility for moving the transaction between the issuing and acquiring banks for every transaction on their network). The issuing bank receives roughly half of the total interchange fee generated on each transaction.

DELINQUENT & CHARGED-OFF ACCOUNTS AND BALANCES

Delinquent accounts are those that fail to pay their minimum payment in any month. Credit issuers track delinquency for up to 6 months past the account's original due date.

There are a variety of patterns to delinquency. Some accounts become delinquent for a single cycle and then pay it off and stay current, and only fall into the delinquent status once or twice a year. Other accounts fall into delinquent status once and stay there for several months, possibly paying off purchases from a holiday gift buying spree, or getting into a situation where they are using their credit card as a form of liquidity for paying off shorter-term debts. Some accounts can become significantly delinquent into multiple months and never regain current status, and eventually the issuer charges off the outstanding balance for this account.

Issuers track delinquency in terms of a metric called "days past due". Typically the cardholder has a full cycle to pay the minimum payment; on the first day of the following cycle, if the payment has not been received, the account becomes one day past due. Accounts that are between 1-30 DPD are usually binned together (called the DPD1 bin) for delinquency treatments, as are accounts between 31-60 DPD, 61-90 DPD, and so on in sets of 30 days past due per bin, up to the typical maximum of 180 DPD. Once an account has gone beyond 180 DPD, the account goes into collections. The naming convention for each bin of delinquent accounts is DPD1, DPD2, DPD3, …, DPD7 (this last bin is the set to be charged off).

RECOVERED BALANCES

As mentioned above, accounts past 180 DPD are sent to collections. Some issuers try to collect balances on some accounts themselves, and refer other accounts to collections agencies. At this point, most issuers focus on the dollars to be recovered, and not the account relationship, so the primary metric is the outstanding balance.

GROSS AND NET LOSSES

Balances that are charged-off by the issuer are called gross credit losses ("GCL"). Some portion of GCL is recovered through collections. What remains is called net credit losses ("NCL"). The ratio of GCL to total outstanding balances is called the GCL rate (and likewise for NCL rate), which are metrics that most credit issuers track actively.

Derived metrics

The section above presents the core metrics that are typically of greatest interest to decision makers because they appear on financial statements and they form the major trends of the bank's performance. However, for analytical reasons, most of the core metrics (with the exception of active accounts) are not the ones that are used to build the statistical forecasts that form the core of the revenue and loss forecast. Instead, we calculate and forecast a number of derived metrics, because a major element of the forecasting process is involved with accounts that are so new to the bank, or have not yet been acquired, that they cannot be statistically forecasted. See the discussion on account lifecycle and vintage maturity in the "Conventions" section below for more details.

SALES VOLUME PER ACTIVE ACCOUNT

Equal to total sales volume divided by total active accounts (per segment, vintage and performance month).

OUTSTANDING BALANCE PER ACTIVE ACCOUNT

Equal to total outstanding balance divided by total active accounts (per segment, vintage and performance month).

FINANCE CHARGES PER ACTIVE ACCOUNT

Equal to total finance charge divided by total active accounts (per segment, vintage and performance month).

LATE FEES PER ACTIVE ACCOUNT

Equal to total late fees divided by total active accounts (per segment, vintage and performance month).

INTERCHANGE PER ACTIVE ACCOUNT

Equal to total interchange fees divided by total active accounts (per segment, vintage and performance month).

PERCENTAGE OF ACCOUNTS THAT ARE DELINQUENT OR CHARGED-OFF

Equal to the sum of accounts that are in the bins DPD1 through DPD7, divided by the total active accounts (per segment, vintage and performance month).

PERCENTAGE OF DELINQUENT ACCOUNTS THAT ARE IN DPD1, DPD2, ... DPD7

This is not one, but seven separate metrics. Each metric is equal to the number of accounts in any one DPD bin, divided by the sum of accounts that are in all bins DPD1 through DPD7 (per segment, vintage and performance month).

PERCENTAGE OF CHARGED-OFF BALANCES THAT ARE RECOVERED THROUGH COLLECTIONS

Equal to the ratio of dollars recovered to the gross charge-off balances.

Example: mapping of core and derived metrics

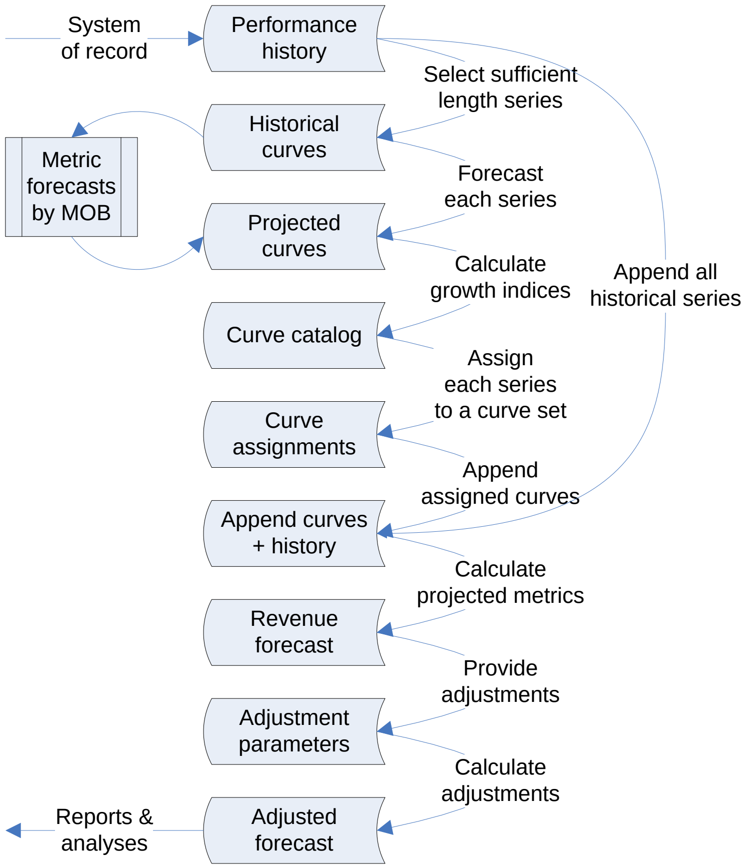
The following table was developed for a proof of concept focusing on a loss forecast, and illustrates the mapping of core metrics to derived metrics. Items highlighted in yellow were used in developing a statistical forecast.

Category	Subject	Accounts	Dollars	Account %	Dollar %	Dollars / Acct
Open	All	open_accounts	tot_os			tot_dpa
New	All	new_accounts				newacct_dpa
DLQ	All	dlq_accts_all		dlq_all_shr		
	DLQ1	dlqaccts1		dlqshr1		dpa1
	DLQ2	dlqaccts2		dlqshr2		dpa2
	DLQ3	dlqaccts3		dlqshr3		dpa3
	DLQ4	dlqaccts4		dlqshr4		dpa4
	DLQ5	dlqaccts5		dlqshr5		dpa5
	DLQ6	dlqaccts6		dlqshr6		dpa6
CO	All	co_accts	tot_cl	co_shr		
	Policy	policy_accts	policy_os	policy_shr		policy_dpa
	BK	bk_accts	bk_os	bk_shr		bk_dpa
	Deceased	dec_accts	dec_os	dec_shr		dec_dpa
Recovery	All		recovered_os		recov_shr	
Reversal	All		reversal_os		rever_shr	
	Fin Chg		finchg_reversal		fcrv_shr	
	Other		(other_reversal)		othrv_shr	

= forecasted metric

Flow chart

The following flow chart illustrates the steps involved in steps 1 and 2 of either the revenue forecast or the loss forecast. The steps are basically the same in each half of the forecasting process, differing essentially on the metrics involved in the specific revenue or loss forecast, and some features of the method of building the statistical forecast.



Advanced topics

Selecting proxies

When confronted with forecasting performance for new and recent vintages, a decision needs to be made with respect to which mature vintage(s) should be used to assign curves to these new and recent vintages. Sometimes, judgment and experience with the performance of past vintages is a good guide, especially if management has well-founded expectations for performance based on the attributes of the new and recent vintages that were targeted at the point of acquisition.

In some instances, no single mature vintage might be the best proxy for a new or recent vintage, and instead the composite curve at a segmentation level above the vintage could be a more desirable choice.

If account-level information is available for recent vintages, this information can be summarized into a profile of the recent vintage and compared with the profiles of mature vintages to select the best proxy. For new vintages, if expectations for the prospecting and booking criteria are available, that could be used to compare with similar criteria that were used with mature vintages when they were originally prospected.

The control tables used in a forecast scenario, containing curve assignments for each of the revenue and loss forecasts, are the tool to use in modifying the proxy assignments. One could create multiple versions of proxy assignments based on alternative criteria, and run multiple forecast scenarios, each of which tests the viability of alternate proxy assignment approaches on the final forecast.

Variance analysis

Once a forecast scenario has been created, deployed and monitored against updated actual performance for the forecasted period, the forecasting analyst should spend time analyzing the variance between actual and predicted performance. The results of the variance analysis can be in one of three categories:

- the forecast is producing satisfactory results compared with actual performance,
- the forecast needs to be tuned or calibrated for the following period given the differences noted in the current actual period,
- the forecast provided apparently valid projections given assumptions used at the time of the forecast publication, but some event or significant adjustment in industry conditions (such as one or more macroeconomic drivers) occurred which led to a large variance in the actual performance compared with the forecast.

Variance analysis can be performed at different points in the forecasting maturity process (refer to the section above titled "Stages of forecast process maturity"). At stage 2 in the process, the variance analysis will likely be conducted with respect to a period of six to twelve months in the performance history. At stage 4, once the forecast has been deployed and is compared with new actual performance results, the variance analysis might instead focus on the most recent performance month, with the intent of improving the following month's forecast. This is called "one-step-ahead" forecast variance analysis. A different type of variance analysis might be conducted as the bank approaches the end of the fiscal year, with respect to end-of-year earnings analysis and business planning for the following fiscal year. In this event, the variance analysis might focus on the fiscal year to date, rather than just the latest performance month.

Forecast adjustments

Let's briefly reflect on the essential basis of statistical forecasting. A statistical forecasting model attempts to project the likely future performance by building a mathematical abstraction of the past performance as captured in the data generated by the underlying business process (i.e., lending people money and expecting them to pay it back, with interest and fees added). Abrupt departures in the way this process will run in the next performance period, especially when this change in the process cannot itself be forecasted statistically (because it is entirely driven by the bank's management). This creates a fundamental challenge for any statistical model, and this is why judgmental adjustments get applied to forecasts, over and above the statistical forecast itself.

Adjustments to the forecast can be executed in a number of places in the forecasting process; nicknames like “front-end”, “back-end” and “top-down” are sometimes used to characterize the step in the forecast process where these adjustments are made.

Front-end adjustments are usually made by the forecasting analyst, likely as a result of variance analysis for quality assurance purposes, prior to broad review of the forecast by management stakeholders. These front-end adjustments are usually made to discrete sets of detailed assumptions, either at the level of a statistical forecasting component, or to a proxy curve assignment decision for a new or recent vintage. If the adjustment is made to a forecast component, this can usually be performed inside the SAS Forecast Studio application, either by editing the model being used for a series, or to insert a judgmental adjustment, using the feature called “overrides”.

Back-end adjustments are typically built to reflect changing management strategies that will have a substantial effect on the nature of expected performance that are not reasonably expected to be captured by the statistical models of past performance. Consider a scenario where senior management decides (as a fairly dramatic departure from past policies) to increase monthly acquisition spending by +75% per month for the next 6 months. By doing so, the expected number of booked accounts in sub-prime credit quality segments will accordingly rise by roughly 40%, because the number of high credit quality prospects is fairly limited. Hence, that decision will likely have a significant impact on the forecast. Some type of back-end adjustment could be applied to the expected delinquency rates for segments where these sub-prime quality accounts end up.

Top-down adjustments typically reflect the presence of a higher-level forecast into which the revenue and loss forecast is expected to contribute and align. A good example of this is the corporate earnings forecast, of which the revenue and loss forecast plays an important role. This earnings forecast is usually owned by the finance function, who typically does not monitor their forecasts at the same level of depth and granularity as the revenue and loss forecast team. A simple way to ensure some alignment is to select a single metric to adjust in the revenue and loss forecast, such as the expected number of active accounts, and then re-run the adjusted forecast accordingly. This adjustment can be made manually, or it could be applied in the active account statistical forecast, utilizing the Forecast Server features of reconciliation as well as overrides (which were designed to work together for this kind of business purpose).

Conclusion

Do you need help with your revenue and loss forecasting processes and practices? Give Robin Way a call at **503.295.1685** or email him at **president@coriosgroup.com**.

About the Author

Robin Way

Robin Way is the Lead Faculty Member for Banking at the International Institute of Analytics, and is the founder and CEO of the management analytics consultancy, Corios. He has over 25 years of experience in the design, development, execution, and improvement of applied analytics models for clients in the credit, payments, lending, brokerage, insurance and energy industries. Robin was previously employed with SAS® Institute's Financial Services Business Unit as a managing analytics consultant for 12 years, in addition to another 10+ years in analytic management roles for several client-side and consulting firms.

Robin's professional passion is devoted to democratizing and demystifying the science of applied analytics. His contributions to the field correspondingly emphasize statistical visualization, analytical data preparation, predictive modeling, time series forecasting, mathematical optimization applied to marketing, and risk management strategies. Robin's undergraduate degree from the University of California at Berkeley and his subsequent graduate-level coursework emphasized the analytical modeling of human and consumer behavior.

About Corios

Corios is a leader in the discipline of management analytics consulting focused on helping clients across industries to bridge the gap between their data and their business decisions. The company's custom offerings provide analytical solutions for risk management and compliance, marketing, pricing, and big data initiatives. Corios' solutions have identified business improvements worth hundreds of millions of dollars for corporations across North America.

Corios believes that data-driven techniques are the key to making profitable business decisions, and that data should be simple, approachable, and implementable for both analysts and decision makers. The Corios team is persistent in their efforts to promote transparency and common process across the analytical discipline.

To learn more about how Corios is bridging the gap between data and business decisions, please visit **coriosgroup.com**.

Revenue and Loss Forecasting
for Banking:
Fundamental Principles of
Corios Forte

Published by Corios
133 SW Second Avenue
Suite 202
Portland, Oregon 97204

CORIOS

WE TELL THE STORY
NUMBERS CAN'T

Visit our website at: **coriosgroup.com**