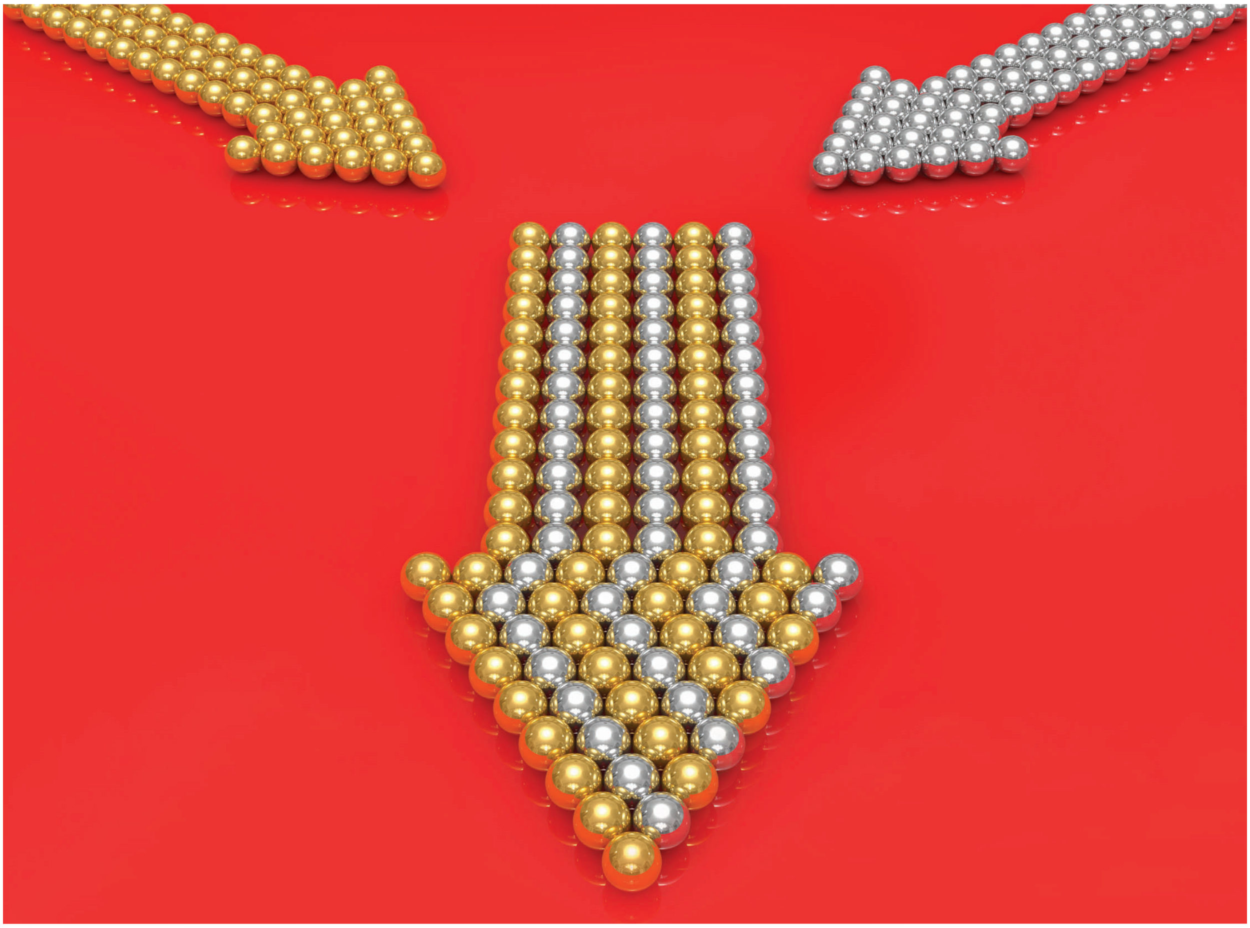
Marrying Credit Scoring and Time-Series Data

••Portfolio and individual loan forecasting can benefit from careful modeling.



Marynchenko Oleksandr/Shutterstock

BY JEFFREY S. MORRISON

WITH TODAY'S ECONOMIC landscape and the advent of Basel II requirements, risk practitioners are struggling to integrate economic time-series data with their existing credit-scoring models to better understand portfolio risk.

Many practitioners see this as a deceptively simple marriage, but the truth may require somewhat of a paradigm shift. Just as hybrid technology is driving change within the transportation industry, the union of credit scoring and time-series data has the potential to promote change inside the risk management arena. This article examines the proposed matrimony on two fronts—where portfolio forecasting is the objective and where the aim is to rank-order credit risk at the account level.

Forecasting Portfolio Performance

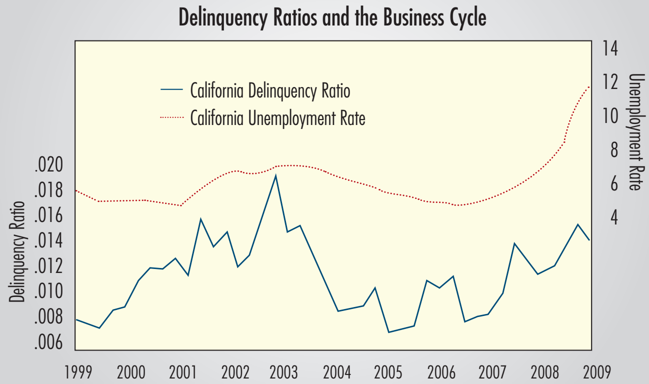
The statistical methodology behind credit scoring has been around for many years. Its roots began as early as 1936, when a statistician named R.A. Fisher used a mathematical approach to identify various species of the iris flower. In 1941, David Durand began formulating applications to loan data using very similar techniques.[1]

Since then, and especially during the 1980s, the use of scoring methods mushroomed as lenders made credit cards almost as ubiquitous as cash. Today, logistic regression for use in probability of default and scoring models has captured the lion's share of the market when it comes to calibration techniques—a methodology whose strength shines when the purpose is to differentiate between good and bad loans at the individual account level.

Although the success of this and similar binary outcome techniques using account-level data is beyond reproach, can the same method be used to forecast performance at the portfolio level as we move through the business cycle? For example, Figure 1 shows that the unemployment rate in California is closely related to bank card delinquency, not only during our present economic crisis, but also historically.

Figure 1

Bank Card Delinquency and Unemployment



Source: Delinquency ratio from trend data of transUnion, unemployment rate from Economy.com

In an effort to meet Basel II regulations, many financial institutions are taking their probability of default (PD) model, likely developed by a regression technique described earlier, and squeezing in economic data to get to portfolio results. For example, if the aggregation linkage is made at the state level, then every individual in the PD sample residing in California is assigned an unemployment rate equal to that for the state overall.

The problem is that the marriage of individual-level credit data to geographic-based data using this type of regression causes statistical abnormalities. The statistical literature refers to these problems as *aggregation bias*. Writes economist Thomas Garrett: "The use of aggregated data to explain individual behavior makes the assumption that the hypothesized relationship between the economic variables in question is homogenous across all individuals. When the behavior of economic agents is not the same, a regression analysis using aggregated data can provide conclusions regarding economic relationships that are different than if less aggregated data were used."[2]

In practice, what this often means is that there is a tendency to find statistical significance in economic relationships when none really exists. Furthermore, using aggregate information within the context of individual-level regression models produces an inflated measure of fit. "Reductions in the amount of variance in the data through aggregation can both eliminate important information and falsely inflate the value of R^2 yielded by regression analysis,"[3] according to a Georgia Tech study. As a result, true forecasting error is adversely affected.

Knowledge of these statistical issues has been well documented for years, so why are these methods still pursued by many risk practitioners? The answer may be twofold: simplicity and inertia. It is a simple matter to pluck a state-level unemployment rate from a data source, assign it to an individual account, and then hit the rerun button on your computer. Furthermore, there is a human tendency to view the familiar as a one-size-fits-all approach, regardless of the ultimate objective. Yet, the Basel II objective is *portfolio* forecasting, not a rank-ordering of individual risk. Therefore, additional approaches are required.

If the portfolio manager wishes to understand the impact of various economic cycles on the book of business, forecasting models are needed that use aggregated data under a regression tool properly suited to that purpose. The recommended approach given this objective is to construct a macroeconomic model using portfolio data aggregated at some geographic level. For example, once the lending institution obtains portfolio measures of its delinquency (loss) rate over the past five or more years by state or metropolitan area, a pooled regression model can be designed to predict loss.[4]

Here we might specify loss rates as a function of portfolio-specific policy variables, some measure of credit scores (mean, median, percentage of accounts by score band), and some pertinent economic time-series (unemployment rate, housing prices, interest rates, and real income per household). Once estimated, projected values of the explanatory variables are required to forecast what may be termed a "base-case forecast" under business-as-usual or most-likely conditions. For the stress test, the analyst can perform "what if" scenarios as needed on the economic variables as a whole or under certain situations, independent of one another.

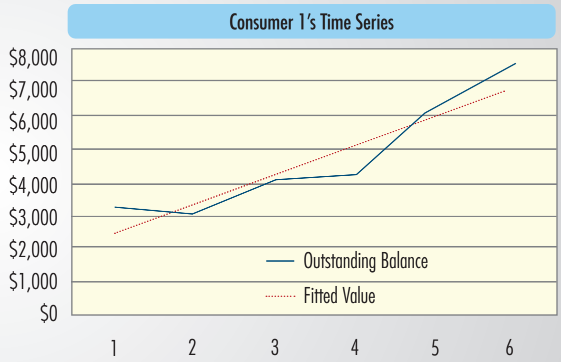
Finally, this type of modeling framework allows us to properly identify lagging and leading relationships in the data, which is more elusive under a strict, traditional credit-scoring approach. For these reasons and others mentioned earlier, if the objective is to paint a portfolio-level picture of future landscapes under changing economic conditions, the

Figure 2

Constructing New Time-Series Trend Characteristics

ID	Time Trend	Outstanding Balance	Fitted Value
Consumer 1	1	$3,200	$2,550
Consumer 1	2	$3,100	$3,440
Consumer 1	3	$4,200	$4,330
Consumer 1	4	$,4,400	$5,220
Consumer 1	5	$6,250	$6,110
Consumer 1	6	$7,500	$7,000

ID	Regression Slope	P-Value	New Trend Characteristic
Consumer 1	890	0.003685	890
Consumer 2	544.2	<.0001	544.2
Consumer 3	72.9	0.5411	0
Consumer 4	-344.4	0.001322	-344.4



Fitted Value = 1660 + (890*Time Trend)

forecasting accuracy from a portfolio-level approach is more accurate than those derived from the bottom up.[5]

Rank-Ordering Account-Level Risk

Historically, the development of scoring characteristics is based on static snapshots of the consumer credit file at a single point in time. However, it is well known that the propensity for debt repayment is sometimes impacted by how recently a behavior occurred—for example, the number of inquiries three months ago or six months ago, or the number of trades opened in the past three months or the past six months. Developing these types of characteristics expands the number of modeling variables from purely static ones to characteristic sets that try to incorporate, to some extent, time-series behavior.

A consumer who has taken steps to improve his or her credit score over the past year might be less of a risk than someone who has done nothing to push the score upward.

The construction of these additional characteristics is easy to develop, requiring the same programming approach needed to produce the other basic characteristics. However, because they have to be entered into the modeling process in a form independent of one another, important information about the impact of historical trends may not be captured adequately. It is this type of information that could enhance the predictive power of the model.

For example, one might expect that a consumer with a downward-trending credit score over the past 12 months should be at greater risk of nonpayment than one who exhibited no trend historically, all other things remaining equal. Conversely, a consumer who has taken steps to improve his or her credit score over the past year might be less of a risk than someone who has done nothing to push the score upward.

This perspective could apply equally well to the fundamental characteristics within the credit score itself. Let's say we know that utilization positively impacts credit risk—something that is built into most credit-scoring models. A customer showing a consistent upward trend in utilization patterns could very well be riskier than a consumer who has the same level of utilization today but shows no trend at all historically.

Now, let's take a look at a different approach to the union of credit scoring and time-series data, keeping in mind that the objective is to enhance the rank-ordering of risk at the account level. This method centers on the construction of a different class of predictive characteristics and incorporates the findings of statistical analysis directly into the variables themselves. However, these new characteristics should always be viewed as a supplement to your existing set of credit-scoring variables, not as a replacement.

In order to create these new characteristics, it is necessary to have consumer-level time-series data on each standard credit characteristic representing a minimum of six consecutive monthly or quarterly snapshots. For illustrative purposes, let's say we have six months of consecutive credit history on each consumer in our sample. The creation of these new characteristics consists simply of running a linear regression on each consumer history. For example, if we are creating a utilization time-series characteristic, we would regress a simple time trend (1, 2, 3, 4, 5, 6) on utilization across the six monthly snapshots of data.

Once the regression is estimated, the value of the slope (trend) and its p-value are used to determine if a trend exists in that characteristic. A p-value above .05 indicates that the trend is no different from zero. As shown in Figure 2, the new time-series characteristic for that traditional score (or variable) would be assigned a value equal to the slope coefficient; otherwise, it would be set equal to zero. Even if the decision is made not to use the new characteristics as part of a formal implementation procedure, such an analysis could yield useful information for understanding the proportion of consumers who have a positive, negative, or no trend in their credit history.

One possible drawback, of course, is the amount of time needed to compute the characteristics. If you have 300,000 consumers in your sample, you have to run 300,000 regressions for each characteristic. But because these are individually very quick processes, the time to create the new characteristics is not as daunting as it might first appear.

This procedure was recently validated at a high level on credit bureau data related to mortgage-backed securities to determine if it added incremental lift over standard credit-scoring approaches. The test consisted of using a *champion* and a *challenger* formulation. For the champion model, a simple logistic regression was estimated using the traditional credit characteristic (utilization, for example) as the single independent predictor to rank-order a sample of good and bad mortgage accounts. A K–S statistic, the classical measure of accuracy, was calculated and recorded. Next, the challenger was estimated using the same champion characteristic along with the new time-series characteristic—for example, the coefficient value (slope) from the regressions.

The results were encouraging. For example, the study showed an incremental lift of more than 80% in accuracy for the time-series characteristic associated with total current balances of all open trades, excluding mortgages. Such evidence supports conventional wisdom that consumers with increasing utilization pose an increased risk over those with flat trends, all other things remaining equal. Therefore, the incremental lift obtained from the new time-series credit characteristics is enough to warrant their inclusion as a *potential candidate* within the traditional credit-scoring model-building framework, implementation issues notwithstanding.

Conclusion

As we have seen, time-series and individual data not only can coexist, but live in harmony. However, a "one size fits all" mindset can limit the effectiveness of any analysis if it is made without regard to the stated objective. The good news is that the risk management profession is already getting better at integrating a number of different time-series techniques into the credit landscape. One such example is survival analysis, which is intended to yield insight into point-in-time default predictions needed for profitability and financial projections.

This article has presented another way to integrate time-series data into the credit risk process—one that deals not with new estimation techniques, but with the design of a new class of predictive characteristics. Although simple and straightforward, this design could potentially improve the development of certain scorecards, especially if they are created or implemented in an environment suffering from a severe economic downturn.

On the portfolio forecasting side, models using panel or pooled data approaches are gaining popularity because they are designed to evaluate changes across the business cycle for purposes of strategic planning and stress testing. These approaches allow for comparisons to be made relative to a base-case outlook and can incorporate a variety of economic scenarios dependent on significant risk factors, such as the impact of government stimulus and assumed trends in oil prices. As financial institutions become more familiar with these techniques, customized software interfaces can be integrated to maximize their usefulness and offer additional transparency to better satisfy regulatory scrutiny. ❖

A "one size fits all" mindset can limit the effectiveness of any analysis if it is made without regard to the stated objective.

Note: TransUnion makes time-series credit data available in an aggregated and depersonalized format only.

***Jeffrey S. Morrison** is senior manager of Research and Econometrics in TransUnion's analytics and decisioning services group. Contact him at* jmorris@transunion.com.

Notes

1. Thomas, Lyn C. 2000. "A Survey of Credit and Behavioural Scoring: Forecasting Financial Risk of Lending to Consumers." *International Journal of Forecasting* 1, no.2 (April–June): 149–72.

2. Garrett, Thomas A. 2002."Aggregated vs. Disaggregated Data in Regression Analysis: Implications for Inference." Federal Reserve Bank of St. Louis Working Paper no. 2002-024B (November).

3. Walker, Neff, and Catrambone, Richard. 1992. "Aggregation Bias and the Use of Regression in Evaluating Models of Human Performance." Georgia Institute of Technology GVU Technical Report no. GIT-GVU-92-07 (March).

4. Morrison, Jeffrey S. 2009."Forecasting Portfolio Performance in an Uncertain Economy." *The RMA Journal* 91, no.10 (July–August): 52–57.

5. Hughes, Tony, and Stewart, Robert J. 2008. "Forecasting and Stress Testing Using Model Pool-Level Data." Economy.com (August).