BEST PRACTICES IN EUROPEAN STRESS TEST MODELING

Dr. Joseph L. Breeden
Chief Executive Officer
Strategic Analytics

3 December 2010

CONTENTS
1. Introduction ........................................................................................................................................2
2. Stress Test Models ..........................................................................................................................3
  2.1. Why retail is different ...............................................................................................................3
  2.2. The Limits of Scores .................................................................................................................4
  2.3. Effective Modeling Techniques ...............................................................................................4
3. Correlating to Macroeconomic Data ...............................................................................................6
  3.1. Transforming Macroeconomic Variables ................................................................................6
4. Designing Macroeconomic Scenarios ............................................................................................7
5. Reverse Stress Testing ....................................................................................................................8
6. Conclusions ......................................................................................................................................8
7. Bibliography ....................................................................................................................................9
1. INTRODUCTION

The recent Global Financial Crisis has highlighted both the need for stress testing and the serious shortcomings in common stress testing approaches. This was observed by the (Basel Committee on Banking Supervision, 2009) and the (Committee of European Banking Supervisors, 2010). Specifically, (Basel Committee on Banking Supervision, 2009) observed:

...given a long period of stability, backward-looking historical information indicated benign conditions so that these models did not pick up the possibility of severe shocks nor the build up of vulnerabilities within the system. Historical statistical relationships, such as correlations, proved to be unreliable once actual events started to unfold.

This observation is telling, because past experience has shown that simplistic models are unable to extract useful information from the mild history that was present in many European countries, but the author also saw that when appropriate models were deployed, the pressures building were visible, reasonable stress tests were possible, and stable correlations could be created. This whitepaper will describe best practices around achieving those goals.

Rather than attempt to summarize the complete stress testing process for ICAAP and European regulators, we will focus on common elements that pose the greatest challenges for retail lenders. Most of the examples provided here come from the US mortgage crisis. Analyses for a number of European lenders have shown similar patterns and conclusions, but their data is proprietary. Therefore, we discuss European results by analogy to this US data.

Creating forecasting or stress testing models for retail lending is different from other bank products, because consumer loans exhibit strong lifecycle effects. New loans and old loans are low risk, but loans two to four years old exhibit significantly higher risk. Therefore, when the industry has an origination boom, such as in 2006-2007 for for many major economies (cite), weaker models can confuse low loss rates due to lifecycle effects with macroeconomic impacts. In addition, detailed analysis of the US mortgage crisis has shown that consumer appetite for loans goes through cycles with interest rates and house prices. This macroeconomic adverse selection means that, although a credit score can rank-order the riskiness of accounts at a specific time, scores alone cannot be used to compare the riskiness of loans booked in one time period to those in previous periods.

For these reasons, simple time series models and credit score models failed during the Global Financial Crisis and are equally weak for the stress testing. This has been proven to be true for all retail loan types, not just mortgage. Stress test models must include both the lifecycle and macroeconomic adverse selection effects so that the sensitivity to key drivers like housing prices and unemployment rates will be reliable.
In this article, we will discuss

- what is needed for a successful stress test model,
- how to incorporate macroeconomic data,
- how to assign probabilities to specific macroeconomic scenarios, and
- how to conduct reverse stress tests.

2. Stress Test Models

Stress testing guidelines do not specify precisely how stress tests should be conducted, but rather detail the goals of the stress tests. This is a sensible approach given the wide range of products within a bank to be stressed and the unique modeling needs of each. Retail loan products (credit cards, auto loans, mortgages, home equity loans, student loans, and personal loans and lines, and deposit accounts) present unique challenges and therefore require a specific class of models that would not generally be found in other parts of the bank. In fact, the methods that work well for retail can be applied to other loan types (Breeden J. L., 2009), but other methods are often available as well.

2.1. Why Retail is Different

Stress testing can be viewed as forecasting with an extreme scenario, but a stress test extrapolates well beyond the range of historical experience. A sound stress test model begins with creating a reliable scenario-based forecasting model, and then builds in robustness when extrapolating beyond the bounds of past observations. Providing the 24 months of quarterly forecasts is inherently a time series problem, and yet the standard time series methods such as ARMA / ARIMA models breakdown with applied unmodified to retail loan portfolios.

When a new loan is created, the consumer’s risk of default exhibits strong lifecycle effects, Figure 2. These lifecycles have significant consequences for portfolio management and modeling. When a large number of loans are booked, as happened prior to the Global Financial Crisis, those young loans have a lower-than-average risk of default, and therefore lower the blended portfolio delinquency and default rates. As they approach peak credit risk in years two

FIGURE 2: THE RISK OF DELINQUENCY VERSUS THE AGE OF THE LOAN FOR A CREDIT CARD AND A 4-YR TERM LOAN (REPRINTED WITH PERMISSION FROM BREEDEN J. L., 2010).

FIGURE 3: BOOMS IN NEW ORIGINATIONS CAUSE PEAKS IN DELINQUENCY A COUPLE YEARS LATER. REPRINTED WITH PERMISSION FROM (BREEDEN J. L., 2010).
through five, the blended delinquency and default rates rise dramatically. These cycles have been observed many times historically, and are one of the major reasons that simple time series models are not successful for retail loan portfolios.

### 2.2. The Limits of Scores

Many forecasting approaches incorporate scores, but the Global Financial Crisis demonstrated the limits of scores. Between 2005 and 2008 when poor quality loans were being booked in large volumes, many portfolio managers reported that the scores for their new loans were the same as previous pools. While often correct, that was not evidence that the new loans were good quality. Rather, it simply revealed the limits of scores.

Credit scores are based upon specific past performance. They cannot see that something new has occurred. When lenders offered new types of loans with easier qualification criteria in markets that were overheating, this had no impact on the previous payment history of the consumer, and therefore the scores did not change even though consumer risk was much greater.

Credit scores see only past behavior, not psychology. Society can be split into those who are naturally conservative and those who are risk takers. A gambler may have an excellent track record, but those numbers alone cannot prove that he was insightful rather than lucky. Similarly, a good credit score cannot distinguish between someone who has been lucky and someone who is fiscally responsible.

The Global Financial Crisis has again made clear that borrowers can either be shopping for a good deal or betting on a good future. Appetite for credit from fiscally responsible borrowers changes through the economic cycle. When interest rates fall and home prices are flat or rising modestly, value shoppers see a buying opportunity and apply for credit. In 2003 and 2004 when the US Mortgage Crisis began, banks booked huge numbers of good quality loans. However, already in 2005 the conservative consumers were pulling out of the market. By 2007, only the gamblers and fiscally misfortunate remained. This change in consumer appetite through time is being called macroeconomic adverse selection, and is described further in (Breeden J. L., 2010). Once again, such effects are not captured in credit scores.

### 2.3. Effective Modeling Techniques

Stress test models must move beyond dependence upon scores alone, or even stressing those scores, since score histories do not capture the effects described above. Similarly, roll rate models that track how accounts move from one delinquency state to the next must go beyond simple extrapolations of those rates.

Bank analysts have access to a class of models that is ideal to the task of forecasting and stress testing. As a group, these are referred to as *nonlinear decomposition* models. In retail lending, the most well known of these are:

- Survival and Proportional Hazard Models (Hosmer & Lemeshow, 1999)
- Panel Data Methods (Wooldridge, 2002)
- Age-Period-Cohort (APC) Models (Mason & Feinberg, 1985), (Glenn, 2005)
- Dual-time Dynamics (Breeden J. L., 2010)
Survival and Proportional Hazards Models were originally designed to capture the lifecycle effects common in retail loan portfolios. Researchers have recently been expanding them to include macroeconomic impacts as needed in stress testing (Malik & Thomas, 2008), (Belotti & Crook, 2008).

Panel data methods from the onset were designed to capture macroeconomic impacts via specific input factors, and have recently been explored to add credit and month-on-book factors in order to capture retail loan portfolio dynamics.

Age-Period-Cohort models were designed from the start to capture lifecycle, environmental, and cohort effects. They were first developed in demography where they were used to study past trends. When applied to retail lending, the primary changes needed are to think of cohort effects as credit quality effects, and to extend the framework for forecasting.

Dual-time Dynamics (DtD) was developed from the start as a scenario-based forecasting method for retail lending. DtD has been in use for over a decade and applied through a number of crises (Breeden, Thomas, & McDonald, 2008), (Breeden & Thomas, 2008), including several banks in the first US SCAP and 2010 European stress test.

**Table 1: Nonlinear Decomposition Algorithm Details as Applied to Retail Lending. Reprinted from (Breeden J. L., 2010).**

<table>
<thead>
<tr>
<th>Method</th>
<th>Granularity</th>
<th>Event type</th>
<th>Lifecycle</th>
<th>Environment</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survival &amp; Proportional Hazards Models</td>
<td>Account-Level</td>
<td>Terminal Events</td>
<td>Nonparametric</td>
<td>Economic Factors</td>
<td>Scores or Scoring Factors</td>
</tr>
<tr>
<td>Panel Methods</td>
<td>Data Account-level</td>
<td>Any Event</td>
<td>Account</td>
<td>Nonparametric</td>
<td>Economic Factors</td>
</tr>
<tr>
<td>Age Period Cohort Models</td>
<td>Vintage-level</td>
<td>Terminal Events</td>
<td>Nonparametric</td>
<td>Nonparametric</td>
<td>Nonparametric</td>
</tr>
<tr>
<td>Dual-time Dynamics</td>
<td>Vintage-level</td>
<td>Any Account or Balance Rate</td>
<td>Nonparametric</td>
<td>Nonparametric</td>
<td>Nonparametric</td>
</tr>
</tbody>
</table>

All of the methods in Table 1 have been explored as possible approaches for stress testing retail loan portfolios. The pros and cons of these are discussed at length in (Breeden J. L., 2010), but they all share a recognition that capturing lifecycle, environmental, and credit quality effects is critical to successful modeling. Strategic Analytics’ LookAhead Software is based upon Dual-time Dynamics, and been used successfully through numerous crises around the world over the last twelve years.

The advantages just described are essential from meeting the requirements of the stress testing guidelines, (Committee of European Banking Supervisors, 2010), where Guideline 14 specifically requires incorporating new originations effects. Such requirements make a compelling case for using “vintage” models, such as those in Table 1.
3. CORRELATING TO MACROECONOMIC DATA

One of the greatest model failures in the Global Financial Crisis was in computing correlation. Whether the correlation in defaults between loans or the correlation of defaults to macroeconomic factors, difficulties in computing correlation were pervasive (Committee of European Banking Supervisors, 2010). For retail lending, the source of that problem is the same as discussed throughout this paper. Using models that do not factor out lifecycle and credit quality effects leads to unstable correlations, as evidenced in Figure 3.

If instead, we employ one of the models in Table 1, the solution is straightforward. We still need as much time history as possible in order to see the impacts of macroeconomic cycles on consumer loans, but those impacts can be isolated and modeled. The main challenge becomes one of choosing the correct variables and transforming them properly.

3.1. TRANSFORMING MACROECONOMIC VARIABLES

The 2009 US SCAP defined specific macroeconomic scenarios for Real GDP, Civilian Unemployment Rate, and House Prices. Those were reasonable variables to incorporate for retail loan portfolios, so one would expect a similar set of variables from regulators and banks when designing future scenarios. The key is to create scenarios for variables that are close to consumer balance sheets.

A stressed scenario will naturally push the constituent variables to extreme levels. For most of today’s portfolios, that means trying to predict the behavior of a retail loan portfolio in a macroeconomic regime that is not present in the historical data. Model extrapolation such as this is an error-prone process. As stated in section 3.3.58 of (Committee of European Banking Supervisors, 2010),

"The assumption of a linear response of the results to stressed parameters may not always hold and it is therefore crucial for an institution to achieve high awareness of non-linear interactions between macro parameters and stressed parameters. For example, it might be that only at a certain level of stress, certain hedging strategies might break down or – on the contrary – come into effect; a subsidiary may also fail to be liquid only at a certain level of stress triggering further repercussion throughout the group."

No model can be perfect at extrapolating beyond the range of observed performance, but analysts can protect themselves from many model breakdowns by carefully considering how the variables are transformed prior to inclusion in the model. This is the problem of nonlinearity described above. For example, Real GNP is quoted in currency. However, most analysts and government reports will focus on the annual percentage change in GNP. Although good for intuitive understanding, percentage changes are poor from a modeling perspective, because they are asymmetric. Something could rise by an unlimited percentage, but only fall by -100%. Although no adverse scenario would consider a 100% decline, nonlinearities arise from this asymmetry even for smaller changes. This problem was solved long ago in equities models by considering log-changes. Taking the log of the ratio of two numbers separated by a 12-month time period produces approximately the same numbers as percentage change for small changes, but handles the nonlinearities correctly for large changes and extreme scenarios.

Table 2 shows suggested transformations for a range of macroeconomic variables. These transformations generally show little improvement in accuracy when modeling the historical data, but can be critical to capturing the correct impacts of extreme scenarios.
TABLE 2: SUGGESTED TRANSFORMS FOR MACROECONOMIC VARIABLES TO CAPTURE NONLINEAR EFFECTS FROM EXTREME SCENARIOS.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Common Approach</th>
<th>Preferred Approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Real GDP, Nonfarm Payroll,</td>
<td>Percentage change</td>
<td>Log-ratio</td>
</tr>
<tr>
<td>House Price Index, Unemployment Rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interest Rates, (any variable between 0</td>
<td>Direct Value</td>
<td>Log value</td>
</tr>
<tr>
<td>and infinity)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployment Rate, (any variable between 0 and 1)</td>
<td>Direct Value</td>
<td>Log-odds</td>
</tr>
</tbody>
</table>

4. DESIGNING MACROECONOMIC SCENARIOS

Designing macroeconomic scenarios presents challenges beyond economics. Economists will undoubtedly start with general guidelines about what constitutes baseline or adverse conditions, and seek to create internally consistent scenarios across a range of variables. However, assessing the probability of occurrence of a given scenario is beyond what economists usually provide.

When considering the probability of a given scenario occurring, analysts often confuse point-in-time (PIT) and through-the-cycle (TTC) scenarios, Figure 4. Because of Basel II, most analysts are familiar with creating through-the-cycle estimates. TTC scenarios are essentially unconditional scenarios. They are designed to describe what the environment could look like in any year. The intent of Basel II was to set aside capital for any year, regardless of the current environment. Basel III takes this a step further and makes adjustments for downturns, but the TTC concept remains at the heart of the calculations.

To assess the probability of occurrence for the macroeconomic scenarios, analysts should not use the TTC distributions of Basel II, but must instead compute the conditional probability of occurrence given today’s current environment. Although an economist may provide an intuitive estimate of such a probability, creating a quantitative estimate of the probability is the natural domain of Monte Carlo simulation.
Using Monte Carlo simulation, an analyst can randomly generate many alternate futures for the environmental impacts. Although setting up Monte Carlo models can be quite involved (Breeden & Ingram, 2009), the final step is simply to compare the scenario created by the economist to the distribution of numerically generated scenarios. Strategic Analytics offers a retail specific Monte Carlo scenario generator as the heart of TrueCapital, which can be used for just this purpose.

5. **REVERSE STRESS TESTING**

One of the newest stress testing requirements is around reverse stress testing (Committee of European Banking Supervisors, 2010), i.e. choose a severe outcome of interest in terms of accounts or balances, and solve for a scenario that would create such a dire outcome. Essentially, one must conduct a search for scenarios that produce the outcome of interest. LookAhead’s Goal Seek Tool does exactly this. The software solves for the environmental impacts required to obtain a specific outcome.

The last question is how likely such a scenario would be. For that, the environmental impacts scenario obtained via the goal seek is compared to the distribution down in Figure 4 to compute with the PIT or TTC probably of occurrence.

Note that we do not need to work with specific macroeconomic variables to achieve the desired result. In fact, the process is more accurate to stay within the nonparametric environment function defined in Dual-time Dynamics. After the net environment function (call the exogenous function in Dual-time Dynamics) is created, we can use a stress test model such as described earlier to find macroeconomic scenarios that would produce the given exogenous function.

6. **CONCLUSIONS**
Stress testing has become an essential tool for portfolio management for bankers and regulators. However, creating reliable stress test models for retail loan portfolios is not a trivial activity. Standard time series methods have proven to be ineffective when they are not designed to capture the basic dynamics of retail loans. Effective methods do exist and have been successful in past crises around the world, but they have not been commonly used in retail lending. If lenders and regulators are to make better policy decisions on the basis of the stress test results, they will need to start with better models.

7. BIBLIOGRAPHY


