A Practical Approach to Building CCAR Loss Forecasting Models in SAS 9.3

Andre Toman
VP, Quantitative Risk Modeling Manager in Credit Risk
Advanced Modeling Team Lead
atom@mtb.com
March 12th, 2014
Disclaimer

Any opinions, advice, statements, or other information or content expressed or made in the following presentation are those of the presenter.
Overview

- Purpose of This Presentation
- Expected Loss Framework
- Model Fitting
- How the Model Works
- Backtesting Results
- CCAR 2015 Scenario Results
- Q&A
Purpose of This Presentation

• Decompose the stress testing modeling process for consumer portfolios into manageable components
• Provide examples that in whole, or in part, may be adapted to current modeling processes resulting in lift
• Provide a foundation of knowledge that can be useful to modeling shops that are beginning to build in-house stress testing solutions
• Motivate those interested in building in-house models, but who currently are beholden to expensive consultants or licensed black box software
• Pack what could have been 30 hours worth of material into 30 minutes the best I can
• Ultimately, give back to the SAS community from whom I was able to learn many skills and techniques that I’ve adopted for the purpose of building stress testing models
EL\(^1\) = Probability of Default x Exposure at Default X Loss Given Default
Model each component separately
Balances, early payoff, prepayment, amortization, involuntary payoff, and the like are also important but not technically risk weight parameters. These will be needed for 9 quarter loss rate calculations.

## Model Fitting

<table>
<thead>
<tr>
<th>Component†</th>
<th>Techniques to Consider</th>
<th>Available Model Fitting Procedures in SAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>GEE, G-Side Random Effects, R-Side Random Effects, cubic splines</td>
<td>GENMOD, GLIMMIX, HPGENSELECT, HPLMIXED, REG, TRANSREG</td>
</tr>
<tr>
<td>EAD</td>
<td>Amortization schedules, cubic splines, credit conversion factors</td>
<td>Data step, REG, TRANSREG, SQL</td>
</tr>
<tr>
<td>LGD</td>
<td>Fractional Logit, Weighted Logistic Regression</td>
<td>GLIMMIX, LOGISTIC, NLMIXED, HPLMIXED, HPLOGISTIC</td>
</tr>
<tr>
<td>Payoff</td>
<td>Same as PD</td>
<td>Same as PD</td>
</tr>
<tr>
<td>Balances</td>
<td>Same as EAD</td>
<td>Same as EAD</td>
</tr>
</tbody>
</table>

†The Basel risk parameters will be the focus of this presentation, although many of the same techniques apply to payoff and balances
Model Fitting – Probability of Default

Step 1. Business Knowledge (e.g., 350 <= FICO <= 850)
PROC FREQ for categorical variables
PROC UNIVARIATE for continuous variables
Identify key fields (balances, FICO,
estimated loan-to-value, interest rates,
line/loan terms, acquired vs. core, 141R/ SOP03-3,
lien status, current delinquency state and
next month’s delinquency state, payoff date,
chargeoff date, etc.)

Step 2. Examine the empirical migration matrix

<table>
<thead>
<tr>
<th>Table of Current Month’s Status by Next Month’s Status</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current Delinquency Status</strong> (Payments Behind)</td>
</tr>
<tr>
<td><strong>Row %</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td><strong>Current</strong></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
</tr>
</tbody>
</table>
Step 3. Independent Variables, create as necessary (lags, log ratios), use one continuous variable in place of class variables or many binary variables when possible.

Unemployment Scenarios CCAR 2014

Utilization by MOR for IOLI CD I
This is the “payment shock” phenomenon that the Fed is so concerned about!
Model Fitting – Probability of Default

Step 4.

Fit and store the model for later use

(R-side random effects model shown)

```sas
proc glimmix data=migration_&mig.;
   class FullAccountnumber Date;
   model transition_cd(ref="&c.")=
       Unemployment_rate
       HybridRescore
       Maturation
       Seasonal
   / dist=binary link=logit s;
   random _residual_/ subject=fullaccountnumber type=cs;
   store build.&lib._rside_&c._&m.;
run;
```

See the following SAS Press books for further information:
1. *Overdispersion Models in SAS* by Morel and Neerchal
Model Fitting – Exposure at Default

Determine the appropriate methodology...

1. If it is installment, arithmetically calculate balance
2. If it is purely revolving, examine the marginal increase in utilization at each delinquency state, perhaps assume line is completely drawn at default, e.g., credit card
3. If it is a combination of both, e.g., home equity line of credit, take a hybrid approach
   a. Treat draw period as revolving as appropriate
   b. Treat repayment period as installment
Notice that the parameter estimates for the Fractional Logit and the Weighted Logistic Regression are the same. See SAS Paper 1304-2014 “Modeling Fractional Outcomes with SAS” by Liu and Xin for more information.
## Andre's Model Fitting Tips

<table>
<thead>
<tr>
<th>Do...</th>
<th>Don't...</th>
</tr>
</thead>
<tbody>
<tr>
<td>... write your own algorithms to test hundreds of thousands of models for model selection to ensure significance, proper sign, out-of-sample predictive power, and ultimately reduce model risk.</td>
<td>... use built-in automatic selection techniques† such as forward, backward, stepwise. I've seen too many times when the coefficients have the &quot;wrong&quot; sign, e.g., negative sign on unemployment to predict missing a payment.</td>
</tr>
<tr>
<td>... keep the transformations meaningful: log ratios of non-stationary variables, lags and log ratios of stationary variables. De-trend with log ratios where appropriate, e.g., use log(HPI/lag#(HPI)) vs. raw HPI or lags of HPI.</td>
<td>... use lags much further back than 3, maybe 6 months. If you use a 39 month lag, the effect will never show up in the stress test!</td>
</tr>
<tr>
<td>... pre-screen macroeconomic variable combinations that exhibit a high degree of collinearity. Changes in HPI, unemployment rate, and 30 year mortgage rate, are highly collinear, not to mention other Fed variables.</td>
<td>... use LOGISTIC for correlated data. You would be breaking all sorts of assumptions, and the models don't perform as well as models fit with GEE in GENMOD or R-side random effects models in GLIMMIX.</td>
</tr>
</tbody>
</table>

† See the following link for a balanced treatment of the benefits and drawbacks of using built in automatic selection techniques: http://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_glmselect_sect019.htm
How the Model Works

<table>
<thead>
<tr>
<th>Delinquency State</th>
<th>State Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>State 0</td>
<td>Current</td>
</tr>
<tr>
<td>State 1</td>
<td>1 to 29 DPD</td>
</tr>
<tr>
<td>State 2</td>
<td>30 to 59 DPD</td>
</tr>
<tr>
<td>State 3</td>
<td>60 to 89 DPD</td>
</tr>
<tr>
<td>State 4</td>
<td>90 to 119 DPD</td>
</tr>
<tr>
<td>State 5</td>
<td>120+ DPD</td>
</tr>
<tr>
<td>State 6</td>
<td>Closed</td>
</tr>
<tr>
<td>State 7</td>
<td>Chargeoff</td>
</tr>
</tbody>
</table>

**Legend**
- [Green] Accrual
- [Yellow] Non-Accrual
- [Red] Absorbing State

**Notes:**
- Account X Charges off
- Account Y Closes
- Account Z Performs

Accounts transition to various states from month to month
We’re not just looking at Account X, Y, or Z… We are looking at all accounts on the books and simultaneously forecast their transition behavior month to month based on loan-level and macroeconomic factors. Transition probabilities are converted to transition states using MCS. The process is repeated many times.
How the Model Works

Start

Accounts one Month prior to Stress Test

Append New Originations

Score Data and Determine Transition Behavior with Fitted Models

Move the clock forward one month and keep a record of transition behavior

Analyse Results and compute summary statistics

Simulation Loops

End
Backtesting Results
Out-of-Sample 24 Month Backtest
†The model shown was not built before CCAR 2015, but the results shown were generated using the model and the CCAR 2015 idiosyncratic scenarios for M&T Bank.
Review

• Purpose of This Presentation
• Expected Loss Framework
• Model Fitting
• How the Model Works
• Backtesting Results
• CCAR 2015 Scenario Results
• Q&A
Special Thanks

- My beautiful wife and daughter for their support
- PhilaSUG for hosting this event and allowing me to present
- SAS for making the tools possible for my modeling projects