Exposure at Default models with and without the Credit Conversion Factor

Edward Tong, Christophe Mues, Iain Brown, Lyn Thomas

edward.n.tong@bankofamerica.com

Credit Scoring and Credit Control XIV Edinburgh, August 26-28 2015
Disclaimer

The opinions expressed in this document are solely the authors’ and do not necessarily reflect those of the Bank of America or any of its subsidiaries.
EAD (Exposure at Default)

- Basel II/III - requirement for Internal Ratings Based (IRB) Advanced approach for calculating minimum capital requirements, CCAR stress testing
- EAD defined as gross exposure in the event of obligor default, typically in 12 months
- This study: EAD for credit cards (revolving exposures)
EAD model approaches

Notation

\[ E(t_d) = \text{EAD} \]
\[ E(t_r) = \text{balance at time } r \]
\[ L(t_r) = \text{limit at time } r \]

Credit conversion factor (Taplin et al. 2007, Jacobs 2010, Qi 2009)

\[
CCF = \frac{E(t_d) - E(t_r)}{L(t_r) - E(t_r)}
\]

\[
\text{EAD} = \text{Current Drawn} + (\text{CCF} \times \text{Current Undrawn})
\]
EAD model approaches (cont’d)

Notation

\( E(t_d) = \text{EAD} \)

\( E(t_r) = \text{balance at time } r \)

\( L(t_r) = \text{limit at time } r \)

Utilization change (Yang & Tkachenko, 2012)

\[
\text{util}_{ch} = \frac{E(t_d) - E(t_r)}{L(t_r)}
\]

Direct EAD based on OLS (Taplin et al., 2007)

Mixture models (Witzany, 2011, Leow & Crook, 2013)
Research objectives

- To directly model the EAD amount
- Compare performance of direct EAD models with known industry models – Credit Conversion Factor (CCF) and Utilization Change
- Consider segmentation by credit usage to combine direct EAD and CCF models
- Model time to default as a predictive variable using weighted PD approach
Data

- UK bank
- Credit card portfolio
- 3 years of data (2001 to 2004)
- All observations are defaulted accounts
- >10k defaults in total sample
Data (cont’d)

- Response variables of EAD, CCF, Utilization Change

- 11 behavioural variables including
  Committed Size / Amount Drawn, Undrawn Amount Drawn, Drawn Percentage (Credit Usage) Time to Default Rating Class Average Days Delinquent Absolute Change in Drawn Amount
Response variables
EAD and CCF
Fitted EAD distribution
Training set
Zero-adjusted gamma distribution

The probability function of the ZAGA is defined by

\[
f_Y(y | \mu, \sigma, \pi) = \begin{cases} 
\pi & \text{if } y = 0 \\
(1 - \pi) \left[ \frac{\frac{1}{y^{\sigma^2}} e^{-y/(\sigma^2 \mu)}}{\Gamma(1/\sigma^2)} \right] & \text{if } y > 0 
\end{cases}
\]

for \(0 \leq y < \infty, 0 < \pi < 1, \mu > 0, \sigma > 0\)

where \(\mu\) denotes mean, \(\sigma\) scale,
\(\pi\) probability of zero EAD

\[
E(Y) = (1 - \pi) \mu \quad \text{and} \quad \text{Var}(Y) = (1 - \pi) \mu^2 \left( \pi + \sigma^2 \right)
\]
Generalized Additive Model for Location, Scale & Shape


- General framework for regression models

- Response variable $y \sim D(y \mid \mu, \sigma, \nu, \tau)$ where $D()$ can be any distribution (over 70 different types including highly skew and kurtotic distributions)

- Mortgage LGD model (Tong et al., 2013)
ZAGA model setup

\[
\log(\mu) = \eta_1 = X_1 \beta_1 + \sum_{j=1}^{J_1} h_{j1}(x_{j1})
\]

\[
\log(\sigma) = \eta_2 = X_2 \beta_2 + \sum_{j=1}^{J_2} h_{j2}(x_{j2})
\]

\[
\text{logit}(\pi) = \eta_3 = X_3 \beta_3 + \sum_{j=1}^{J_3} h_{j3}(x_{j3})
\]

where \(X_k \beta_k\) denote parametric linear terms, \(h_{jk}(x_{jk})\) denote additive smooths.
ZAGA model development

- Separate model components estimated for $\mu$, $\sigma$ and $\pi$ components
- Developed with stepwise selection based on Akaike Information Criteria (AIC)
- Continuous variables fitted with smoothers based on penalized $B$-splines (Eilers & Marx, 1996)
Results
Mean of non-zero EAD
Results
Dispersion of non-zero EAD

Dispersion increases with Undrawn (%), decreases with credit usage
Benchmark models:
CCF, Util. Change, PD weighted

- 3 CCF models based on OLS, Fractional Response Regression (Papke & Woolridge, 1996) and Tobit models (Tobin, 1956)
- Utilization Change based on Tobit model
- Survival EAD model - PD weighted approach to model time to default using Cox PH regression:

\[
EAD = \sum_{t=1}^{12} \left( \frac{S(t - 1) - S(t)}{1 - S(t = 12)} \times EAD(t) \right)
\]
Benchmark models: Segmentation by credit usage

- Segmentation using combined Direct EAD (ZAGA) and CCF models
- CCF models perform better for low credit usage, Direct EAD model better for high credit usage
- Credit usage of 90% was optimal cut-off for segmentation based on discrimination and calibration performance measures
Validation methodology

- 10-fold cross validation (CV)
- Spearman $\rho$
- MAE from EAD
- $\text{MAE}_{\text{norm}}$ from $\frac{\text{EAD}}{\text{Commitment Size}}$
- RMSE, $\text{RMSE}_{\text{norm}}$
Validation with 10 fold CV

<table>
<thead>
<tr>
<th></th>
<th>OLS-CCF</th>
<th>Tobit-CCF</th>
<th>FRR-CCF</th>
<th>Tobit-UTIL</th>
<th>OLS-EAD</th>
<th>ZAGA-EAD</th>
<th>OLS-USE</th>
<th>ZAGA-USE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spearman ρ</td>
<td>0.741</td>
<td>0.737</td>
<td>0.743</td>
<td>0.746</td>
<td>0.744</td>
<td>0.742</td>
<td>0.750</td>
<td>0.749</td>
</tr>
<tr>
<td>MAE</td>
<td>859.0</td>
<td>870.6</td>
<td>856.1</td>
<td>925.2</td>
<td>883.3</td>
<td>833.5</td>
<td>829.9</td>
<td>819.2</td>
</tr>
<tr>
<td>RMSE</td>
<td>1614.8</td>
<td>1586.3</td>
<td>1577.7</td>
<td>1654.3</td>
<td>1546.1</td>
<td>1602.5</td>
<td>1565.7</td>
<td>1571.0</td>
</tr>
<tr>
<td>MAE&lt;sub&gt;norm&lt;/sub&gt;</td>
<td>0.273</td>
<td>0.276</td>
<td>0.273</td>
<td>0.294</td>
<td>0.301</td>
<td>0.268</td>
<td>0.269</td>
<td>0.260</td>
</tr>
<tr>
<td>RMSE&lt;sub&gt;norm&lt;/sub&gt;</td>
<td>0.432</td>
<td>0.430</td>
<td>0.430</td>
<td>0.442</td>
<td>0.448</td>
<td>0.454</td>
<td>0.430</td>
<td>0.429</td>
</tr>
</tbody>
</table>

ZAGA-EAD, ZAGA-USE have lowest MAE

Survival EAD MAE=830.3, MAE<sub>norm</sub>=0.266
Validation
Observed vs Fitted EAD Densities
Conclusions

- Modelling the EAD amount directly can produce competitively predictive EAD models
- Segmentation by credit utilization offers greater performance benefits - a combined approach with EAD and CCF models may work better
- The time to default variable can be used a priori within a PD weighted model
References


Leow, M., Crook, J. (2013). A Two Stage Mixture Model for Predicting EAD. Credit Scoring & Credit Control XIII. Edinburgh, UK.


