DFAST Credit Loss Modeling with Call Report Data

Clayton Botkin and Jim Van Osten

Many banks are currently working through the challenge of the DFAST stress testing submissions. Integral to this is an analysis is a modeling forecast development for each material asset class in a bank's portfolio. Credit risk modeling for retail and wholesale assets involves not only statistical expertise, but also business acumen and a true sense for what is acceptable in the current regulatory environment.

In this article, we present our development work for a baseline set of DFAST models useful for any bank as a first step in modeling or as a supporting benchmark solution.

As a leader in Model Risk Management, Montana Analytics has been active in developing quantitative models and utilizing rigorous analytical methods for examining models since 2002.

Background and Purpose

Montana Analytics developed proprietary models using Call Report data to generate DFAST forecast losses for a variety of loan products in both wholesale and retail asset classes. The purpose of the model suite is the generation of scenario-specific loss projections as part the DFAST stress-testing exercise. The modeling dependent variable is net charge-off rate and the model suite is known as MADCLM ("Montana Analytics DFAST Credit Loss Models").

This modeling suite can be deployed to fulfill a bank's forecasting needs or separately as an independent benchmark for comparison to internal models. Further, the models can be re-estimated on either peer group or regionalized data to produce more tailored solutions.

Montana Analytics followed a standardized and well-defined approach to develop and test "top-down" credit loss models for use in the Dodd-Frank Act Stress Test ("DFAST") exercise. This document summarizes the approach and provides details on the model development, theory, and results. The models are characterized as "top-down" because they use aggregated bank Call Report data as the source of information for the dependent variables where each model corresponds to specific items in the Call Report. There is a suite of models for many of the major loan categories to forecast losses directly with a single equation for each category – no segmentation exists within each model segment. To illustrate, the "CRE" model is used to predict losses for non-owner occupied Commercial Real Estate ("CRE") loans. This is identified in the Call Report as Item 9 "Non-farm, non-residential other loans"[1].

National macroeconomic data for the supervisory variables in the DFAST scenarios serve as the independent variables. The approach aligns with regulatory and market practices regarding model development and testing for similar DFAST benchmark models at many banks with assets between $10 billion and $50 billion.

Modeling Approach

Montana Analytics used a top-down approach built on a cross-section of national bank data to model net charge-offs. Montana Analytics chose the top-down approach due to a lack of extensive loan-level modeling data to construct models that are more granular. This approach is similar to existing champion models deployed at many banks with assets between $10 billion and $50 billion. The data for each loan asset portfolio are constructed as a time series of varying length.

This methodology is well suited for use in the modeling of data reported at a portfolio level. The approach, utilizing data compiled from numerous banks across the nation, leads to some assumptions that are discussed in the next section. Notably, one bank's data and historical loss experience may differ markedly from those in the development data depending on a range of factors. However, this approach and development data serve to produce useful benchmark models.

Montana Analytics followed a standard modeling approach using Ordinary-Least-Squares ("OLS") regression that is consistent with market practices to develop "top-down" credit loss models for use in the Dodd-Frank Act Stress Test ("DFAST") exercise. The models are characterized as "top-down" as they use aggregated bank Call Report data as the source of information for the dependent variables, where each model corresponds to various items in the Call Report. Call Report data across time for a large cross-section of banks served as the dependent variable, net charge-offs, for each loan segment. National macroeconomic data for the supervisory variables in the DFAST scenarios serve as the independent variables. The approach aligns with regulatory and market practices regarding model development and testing for similar DFAST benchmark models at banks with assets between $10 billion and $50 billion.

Model Assumptions and Limitations

Stress-test models attempt to describe the impact of hypothetical scenarios on real-world outcomes using mathematical constructs. As abstractions of reality, the models rely upon a number of assumptions in order to be tractable, which introduces limitations. Awareness of these assumptions and limitations helps limit the risk that the models are used for purposes for which they are not suited.

[1] This model corresponds to elements RIADC897 (charge-offs) less RIADC898 (recoveries) in the Call Report and to line item 9 on the 2017 FR_Y DFAST form.

Key Assumptions:

1. The model is a simplification of economic relationships. The credit models are developed based on the assumption that the loans in the modeled portfolios are homogeneous with respect to the influence of key risk drivers. Simplification of accounting relationships is similarly necessary to make estimation possible at the portfolio level. Rather than applying Generally Accepted Accounting Principles ("GAAP") standards for charge-offs to individual loans, loan charge-offs are grouped in the aggregate as per the Call Report.

2. Historical data relationships will be maintained in the future. Key empirical relationships from the past captured in the modeling data are robust, implying such relationships will hold in the future across different economic environments

3. There are no significant measurement errors in the model or any of its variables. This assumes that the Call Report data (used as the dependent variables) and macroeconomic data (used as the independent variables) are recorded accurately, and that the explanatory economic variables are correct representations of U.S. government data. Call report data is known to be a fairly reliable source of net charge-offs as it's used by numerous banks for this purpose. It's assumed there isn't a strong systematic bias across the thousands of firms used in developing the models. Said another way, there should not be any strong measurement error for the dependent variables. The sources for the macroeconomic data is the same sources as used by the financial regulators and other market participants, so these models don't introduce any additional measurement error for the independent variables compared to other market participants using similar model specifications.

4. There are no important omitted variables. The macroeconomic variables that comprise each model are the primary drivers of net charge-offs. Though there are no asset-specific variables modeled, this is consistent with the Champion model specifications employed by many mid-sized banks for the DFAST exercise.

5. Estimating net charge-offs involves accounting policy decisions, timing and unknown recoveries. Any action or behavior affecting charge-offs is assumed to be captured by the data and models.

Key Limitations:

1. The models are only useful to estimate the specific asset class for which they were designed.

2. The models are developed on national-level, multi-bank data and results may not align well with a unique bank portfolio or regional footprint.

Modeling Data and Analysis

All commercial banks are required to report data on loan balances, gross charge-offs and recoveries by asset class to the Federal Reserve using the “Call Report”. The call report data used is therefore extensive and includes 1.5 million quarterly observations of over 15,000 commercial banks operating from January 1980 through Q3 2016. This thirty-six-year period includes five recessions that started in 1980, 1981, 1990, 2001, and 2007. The economic conditions over this time frame include numerous dynamics and swings in macroeconomic factors. While the call reports do not include credit quality details for the loans in portfolios, this rich data still provides a means to analyze portfolio-level risk and charge-offs through stress scenario modeling.

Data

Montana Analytics used public data from quarterly financial statements (Call Reports) of the U.S. banks to construct the dependent variable for each model. The final data was created through a joint process of sourcing from both the Chicago Federal Reserve and the FFIEC. We source this first from the Federal Reserve Bank of Chicago Commercial Bank data to obtain history from 1976 to 2010. Then we obtain quarterly updates through 2016 from the FFIEC Central Data Repository's Public Data Distribution site.

The overall data universe from 1980Q1 to 2016Q3 totals 1,587,136 observations. We restricted our analysis to institutions with assets between $1 billion and $50 billion to align our analysis with those regional banks subject to stress-testing requirements for DFAST. We discard all banks with headquarters in the U.S. territories, uninsured institutions, trust companies, and U.S. branches of foreign banks. The resulting observations remaining total 658,947.

The final sample comprises an unbalanced panel of thousands of banks, of which approximately 6,800 were still active as of 2016Q3, down from the 1983 peak of nearly 16,000 banks. The unbalanced nature of the panel data is important. Inclusion of banks that dropped out due to failure or mergers ensures that we did not introduce survivorship bias into the data. The final sample comprises 658,947 records by bank and Call Report date across all key variables.

This Call Report modeling dataset has vast potential for developing national models, as we present here, and also for use in ‘peer group’ and custom ‘regional analysis’. Custom peer groups can be formed through a variety of methods and controls and then the sub-sample data extracted to provide more direct comparison to a bank’s performance. Likewise, regional analysis can be established using specific bank data in a sample created from known regional footprints in specific asset classes. This Call Report modeling dataset enables both techniques.

Independent Variables

We considered the sixteen (16) domestic national macroeconomic variables directly from the Federal Reserve for the 2017 DFAST/CCAR exercise along with various transformations as the primary independent variables for the models. We then computed numerous transformations to enhance our data exploratory analysis. Transformations included quarterly lags ranging from 1 to 4 quarters, quarterly growth rates, and yearly growth rates. More than 100 variables were generated from the core variables and considered for inclusion into the models. Additionally, a one-quarter lag of the dependent variable was tested as another independent variable as detailed in the *Lagged-Dependent Variable (LDV) Analysis* section.

Lagged-Dependent Variable (LDV) Analysis

We examined a basic model estimation to determine if the dependent series is non-stationary by estimating an ARIMA(1,0,0) model, also known as AR(1), with no exogenous variables. This model form simply estimates the dependent variable using a one-quarter lag of the dependent variable as the sole independent variable. If the standard statistical tests fail to reject the null hypothesis that the dependent series is non-stationary we consider the inclusion of a Lagged-Dependent Variable ("LDV") in the model. For each asset class, we explored model forms with and without the LDV term.

Asset-Specific Data

For each of the following asset types we develop sub-samples where asset-detail exists. For example, for Commercial Real Estate (CRE), we examine the records for loan balances of CRE assets and then select all records from the data. This creates asset-specific sub-samples ready for analysis. This is done for each asset-class:

Table 1: Distribution of Loan Observations and History by Asset Class

Asset Class	Code	Observations	Quarters	Started
Closed-end 1-4 Family Residential 1st liens	RS1	48,522	103	3/31/1991
Closed-end 1-4 Family Residential 2nd Liens	RS2	46,159	103	3/31/1991
Closed-end 1-4 Family Residential 1st+2nd Liens	RES	48,586	103	3/31/1991
Revolving 1-4 Family Residential loans	HEL	49,522	116	12/31/1987
Residential and Other Construction loans	ADC	21,355	39	3/31/2007
Residential Construction loans	RSC	20,089	39	3/31/2007
Other Construction loans	OSC	21,163	39	3/31/2007
Multifamily Residential loans	MFR	59,167	146	6/30/1980
Commercial Real Estate, Non-Owner Occupied	CRE	21,490	39	3/31/2007
Commercial Real Estate, Owner Occupied	CRO	21,114	39	3/31/2007
Commercial and Industrial loans	CNI	63,628	131	3/31/1984
Credit Card	CCD	12,397	63	3/31/2001
Auto Loan	AUT	11,998	23	3/31/2011
Other Consumer loans	OTC	29,253	63	3/31/2001
Farm Land loans	FLD	44,919	146	6/30/1980
Agricultural Production Loans	AGP	40,359	146	6/30/1980
Depository	DEP	5,472	63	3/31/2001
Leases	LES	33,508	146	6/30/1980
Non-Depository Finance	NDF	4,327	27	3/31/2010
Other (composite)	OTH	55,588	146	6/30/1980

Source: FFIEC Call Report data, 1980 to 2016Q3

Table 2: Distribution of Loan Balances by Asset Class

Asset Class	Code	2006 Q4 Balance	Percent of Total	2016 Q4 Balance	Percent of Total
		($ Billions)	*percent*	*($ Billions)*	*percent*
Closed-end 1-4 Family Residential 1st liens	RS1	$297.2	21.57%	$485.6	15.21%
Closed-end 1-4 Family Residential 2nd Liens	RS2	$52.5	3.81%	$16.7	0.52%
Closed-end 1-4 Family Residential 1st+2nd Liens	RES	$349.7	25.39%	$502.3	15.73%
Revolving 1-4 Family Residential loans	HEL	$91.7	6.65%	$91.7	2.87%
Residential and Other Construction loans	ADC	$0.0	0.00%	$131.3	4.11%
Residential Construction loans	RSC	$0.0	0.00%	$32.3	1.01%
Other Construction loans	OSC	$0.0	0.00%	$99.0	3.10%
Multifamily Residential loans	MFR	$58.6	4.25%	$184.2	5.77%
Commercial Real Estate, Non-Owner Occupied	CRE	$0.0	0.00%	$368.5	11.54%
Commercial Real Estate, Owner Occupied	CRO	$0.0	0.00%	$210.0	6.58%
Commercial and Industrial loans	CNI	$269.2	19.54%	$410.7	12.87%
Credit Card	CCD	$81.2	5.89%	$117.6	3.68%
Auto Loan	AUT	$0.0	0.00%	$55.1	1.73%
Other Consumer loans	OTC	$6.9	0.50%	$63.2	1.98%
Farm Land loans	FLD	$11.7	0.85%	$30.0	0.94%
Agricultural Production Loans	AGP	$9.8	0.71%	$23.8	0.75%
Depository	DEP	$8.9	0.64%	$7.5	0.23%
Leases	LES	$21.8	1.58%	$23.0	0.72%
Non-Depository Finance	NDF	$0.0	0.00%	$42.7	1.34%
Other (composite)	OTH	$118.5	8.60%	$297.2	9.31%

Source: FFIEC Call Report data, 1980 to 2016Q3

Data Calculations

For each asset-specific sample, we compute a number of important new fields. First, we remove any loss share amount from its corresponding loan balance.

For each bank, we compute:

- Asset-level Loan percentage (%) as Asset Loan Balance / Total Loan Balance
- Non-performing loan balance as the sum of 90 DPD and Non-accrual loan balances
- Net Charge-Off as Charge-off balances less Recoveries

For each asset within each bank we then compute percentages of the following based on the asset-level loan balances:

- 30-89 DPD %
- 90 DPD %
- Non-Accrual %
- Non-Performing %
- Charge-off %
- Recovery %
- Net Charge-Off %

We compute the net charge-off percentage in the current period using the Asset Loan Balance$_{(t-1)}$ in the prior period.

- NCO percentage (%) as Net Charge-Off$_{(t)}$ / Asset Loan Balance$_{(t-1)}$

This is consistent with industry practice and takes into account the delay in recording charge-offs relative to balances. Finally, we annualize this percentage by multiple the quarterly results by 4. For each asset class, the key dependent variable in this analysis is the Net Charge-Off % (e.g. CRE_NCO_pctTAnn).

Table 3: Summary of Annual Net Charge-Off Rates by Asset Class

Asset Class	Code	Average Annual NCO	Minimum Annual NCO	Maximum Annual NCO
		percent	*percent*	*percent*
Closed-end 1-4 Family Residential 1st liens	RS1	0.16%	0.00%	0.91%
Closed-end 1-4 Family Residential 2nd Liens	RS2	0.73%	0.00%	3.57%
Closed-end 1-4 Family Residential 1st+2nd Liens	RES	0.20%	0.00%	1.15%
Revolving 1-4 Family Residential loans	HEL	0.35%	0.00%	1.76%
Residential and Other Construction loans	ADC	1.20%	-0.04%	3.53%
Residential Construction loans	RSC	1.55%	-0.04%	4.51%
Other Construction loans	OSC	1.10%	-0.05%	3.29%
Multifamily Residential loans	MFR	0.14%	0.00%	0.91%
Commercial Real Estate, Non-Owner Occupied	CRE	0.27%	0.00%	0.91%
Commercial Real Estate, Owner Occupied	CRO	0.20%	0.02%	0.49%
Commercial and Industrial loans	CNI	0.51%	0.08%	1.62%
Credit Card	CCD	3.55%	1.93%	7.15%
Auto Loan	AUT	0.32%	0.28%	0.41%
Other Consumer loans	OTC	0.33%	0.00%	1.59%
Farm Land loans	FLD	0.04%	-0.02%	0.25%
Agricultural Production Loans	AGP	0.80%	0.00%	5.28%
Depository	DEP	NA	NA	NA
Leases	LES	0.27%	0.00%	0.81%
Non-Depository Finance	NDF	NA	0.21%	NA
Other (composite)	OTH	1.20%	0.00%	5.30%

Source: FFIEC Call Report data, 1980 to 2016Q3

Modeling Approach and Selection

Montana Analytics followed a model development approach consistent with typical practices to develop and test the macroeconomic linkages for the models. The model-development strategy emphasized parsimonious models while still attempting to capture the important economic sensitivities for each asset class.

Several models were identified as best final candidates, which included different combinations of the national macroeconomic variables and possibly an LDV term. the general goal was to keep the number of independent variables in each model to as few as possible while also choosing models that displayed robustness across the stress scenarios.

We considered the 16 domestic macroeconomic variables directly from the Federal Reserve for the 2017 DFAST/CCAR exercise plus various transformations of these variables for the models as discussed previously in the *Independent Variables* section. From this, we use several statistical methods to examine correlations, clusters of variables and other steps to narrow down the most important macroeconomic variables.

The modeling approach included these components:

- Correlation analysis of the macroeconomic variables
- Cluster analysis to reduce the most significant variables to key groups
- Forward and backward stepwise regression algorithm
- Statistical significance and sign of the coefficient estimates
- Economic rationale for the included variables
- Adjusted R-squared, AIC, in-sample RMSE metrics
- Back-testing using in-sample forecasting performance metrics
- DFAST Scenario forecasts
- Model assumption diagnostics including serial-correlation, multicollinearity, and heteroscedasticity test results to assess validity of model predictability

Modeling Process

1. The modeling process commenced with a correlation analysis of the dependent and independent variables to identify the more significant variables. Variables with a correlation value greater than 0.30 were kept for consideration in the cluster analysis stage of the model development process.

2. A cluster analysis was then used to identify key groups from the correlation analysis to narrow down the set of independent variables considered for the models. Steps one and two produce a narrowed set of potential modeling variables.

3. The potential modeling variables were entered into forward and backward stepwise OLS regression algorithms, which generate the best model from each process based on the lowest AIC value. More detail on these algorithms is described in the *Regression Modeling Approach* section.

4. Models were also estimated where a first-order autoregressive AR(1) term was also considered as an independent variable.

5. A maximum of four final candidate models were generated based on the above two steps: one model from each of the forward and backward stepwise processes that included an AR(1) term as independent variable, and two models that didn't include the AR(1) term.

6. The modeling process generated numerous statistical assumption diagnostics used to examine the validity of model assumptions for all candidate models. Key tests included:

   - Assess multicollinearity: Inspected the maximum Variance Inflation Factors (VIFs) to evaluate and exclude highly correlated variables.

   - Assess Autocorrelation: We performed the Breusch-Godfrey ("BG") and Durbin-Watson tests for on the model residuals to determine if serial autocorrelation is present.

- Assess Stationarity: We performed the Augmented Dickey-Fuller (ADF) test to determine whether the model was stationary. We also generated and examined the ACF and PACF residual plots for approximately 14 lags of the models.

- Assess heteroscedasticity: Examined if residuals variance is constant using the Breusch-Pagan heteroscedasticity test.

- Assess normality of residuals: Determine if residuals are generally normally distributed through the Shapiro-Wilk Normality test and inspection of the model residuals on a Q-Q plot, aka normal probability plot.

- Assess Model linearity: Assessed if the functional form was correctly specified as a linear model using the RESET test and inspection of the Residuals versus fitted values plots, aka scatter plot.

- Outlier detection: Generated the Cook's D test, Bonferroni Outlier test, and inspected the Residual scatter plot for outliers that could greatly influence the model fit.

- The p-value target for all hypothesis tests is 0.05 (i.e., 5% significance level) with a 0.10 limit (10% significance level), except for the Breusch-Pagan heteroscedasticity test, where we use a cutoff of 0.2 (20% significance level).

7. Assess macroeconomic drivers: Select macroeconomic variables for inclusion in model based on business and economic reasoning. Variables with low explanatory power are systematically eliminated. The vendor considered coefficient estimates with p-values above 0.10 statistically insignificant.

8. We performed the following analyses on the DFAST forecasts for the candidate models:

    - Assess Scenario performance: Assess the model predictions follow a logical path in rank ordered scenarios.

    - Assess Sensitivity testing: Evaluate models based on variation in the independent macroeconomic variables.

9. Each of the models resulting from the automated modeling process is then assessed in terms of standard statistical measures, including coefficient signs, residual diagnostics, inspection of autocorrelation function ("ACF") and partial autocorrelation function ("PACF") charts, in-sample backtesting and predictions. Overall predictive power is assessed based on R-squared metrics with model selection also based on AIC.

10. All final models used in the 2017 benchmarking analyses for clients had acceptable results for backtesting analysis and produced appropriate DFAST forecasts with sensitivity to the macroeconomic variables in each regulatory scenario. No models had violations of modeling assumptions that affected the model's predictability based on the diagnostic testing performed.

Regression Modeling Approach

We considered the candidate set of models using both a standard forward and backward stepwise regression selection algorithm, where again all models are OLS regression form. The model fit criterion for the best model chosen from each stepwise selection algorithm (i.e. forward, backward) was based on the lowest Akaike Information Criterion (AIC) value:

$$AIC = 2k - 2 \ln L$$

where L is the maximum likelihood, k is the number of model parameters, and n is the number of observations.

AIC rewards goodness of fit but it also includes a penalty that discourages model overfitting as increasing the number of parameters in the model generally improves a model's goodness of the fit. The AIC statistic is effectively a tradeoff between an improved fit versus the inclusion of additional parameters and is a common output metric to measure this tradeoff for model selection among a number of different specification. This process mitigates the potential for over-fitting models, which is a best practice consideration given the limited time series for some asset classes.

The Forward stepwise regression algorithm implemented is a standard approach. The process involves starting with no variables in the model, testing the addition of each variable using the chosen model fit criterion (AIC), then iteratively adding a variable whose inclusion gives the most statistically significant improvement of the fit. Again, this statistically significant improvement is based on the AIC value. This process repeats until no variable improves the model fit as measured by a lower AIC value.

The Backward stepwise regression algorithm likewise followed the typical approach. This process involves starting with all candidate variables identified after the correlation and cluster analyses, testing the removal of each variable using the chosen model fit criterion (AIC), then deleting the variable whose removal provides the most statistically insignificant deterioration of the AIC value. This process repeats until no further variables can be deleted without a statistically significant loss of fit as measured by a lower AIC value.

References and Citations

Kapinos, Pavel S. and Mitnik, Oscar A., *A Top-down Approach to Stress-testing Banks*, FDIC Working Paper Series, March 2015, FDIC CFR WP 2015-02.

Hirtle, Beverly, Kovner, Anna, Vickery, James and Bhanot Meru, *Assessing Financial Stability: The Capital and Loss Assessment under Stress Scenarios (CLASS) Model*, Federal Reserve Bank of New York Staff Reports, no. 663, February 2014; revised October 2014.

Appendix: Model Specifications

The following pages list various information on the model estimation process including variable names, transformations, presence of lags and statistical significance.

Macroeconomic variables modeled without any transformation have the word 'level' attached to the end of it. Those that calculate the percentage change from a prior period, either quarterly or yearly, have the word 'growth' in it. QoQ represents Quarter-over-Quarter, YoY is Year-over-Year, and log is the natural logarithm. Those variables with a lag have the word 'lag' with a number from 1-4 (representing the number of lags) following it. In addition to the 16 domestic macroeconomic variables that may contain some transformation or lag, Montana Analytics defined other variables from this pool.

These defined variables are:

- BBB Spread = BBB Corporate Yield – X10 Year Treasury Yield
- Term Spread = X10 Year Treasury Yield – X3 Month Treasury Rate
- HPI Max = if House Price Index Level growth1 > 0 then House Price Index Level growth1 else 0

Statistical Significance of model variables is indicated as follows:

Symbol	Significance Level	Confidence Level
***	0.001	99.9%
**	0.01	99.0%
*	0.05	95.0%
.	0.1	90.0%
'	1.0	0.0%

