Z-CONCEPTS

LOAN DEFAULT ANALYSIS: A CASE STUDY FOR CECL

by Guo Chen, PhD, Director, Quantitative Research, ZM Financial Systems

THE DATA

Data Overview

Since the financial crisis banks have been increasingly required by various regulatory bodies to perform more analysis around capital adequacy in the event of another financial downturn. Simply, do banks have enough capital to withstand investment and loan defaults? Starting with the Dodd-Frank Stress Tests and now, more recently, the Current Expected Credit Loss, banks need to forecast asset defaults under various economic scenarios to see if they are adequately capitalized.

To aide in this endeavor, ZM Financial Systems (ZMFS) has performed a detailed analysis of a multi-billion-dollar bank's loan portfolio in an attempt to derive a statistical model for forecasting loan defaults. The bank provided its monthly loan performance data between early 2002 and May 2017. There are over 1.3 million observations from about 40k+ individual loans. The whole data set covers various loan profiles, including:

- Commercial loans (~40%);
- Real estate loans (~15%);
- Construction loans (~12%);
- Consumer loans (~8%); and
- Home equity lines (~6%).

The bank has an internal rating assigned to most loans at origination, and the rating is updated during the lifespan of the loan based on its performance. The bank has developed a formula that gives a clear definition of a default event. This formula uses the loan's internal rating and its accrual status as inputs. For this project, we set out to perform detailed loan analysis, which included assessing the data quality, choosing the appropriate segmentation, calculating the historical default rates, and building dynamic statistic models to establish meaningful forecast of default rate.

Challenges in Dealing with the Data

When preparing for this kind of default analysis, we usually look for access to common loan credit indicators, such as debt service coverage ratio (DSCR), loan-to-value (LTV), etc. It would be even better if these data sets were updated routinely. Since the bank cannot currently provide these data sets, we must solely rely on the assigned internal rating.

Though there are many fields included in the data set, the data is far from complete. For example, the geographical information fields are accessible, but more than 30% of the loans do not have their STATE field populated with a valid value.

Some data provides conflicting information. For example, on certain transactions the system will input the customer account number in the TRREFF field for easier referencing on the General Ledger. Due to human error, we sometimes found the numbers do not match. After discussing the discrepancies with the bank, we chose to remove such records when the amount of such occurrences is insignificant.

There are also certain complicated situations where the bank works with a client to restructure a loan as it enters the default stage. This type of event could affect the reported loan balance significantly, but it cannot really be captured just by checking the loan performance records in the data set. Also, the distribution of loan original amount (ORGAMT) is very skewed, as can be seen in Figure 1: *Distribution of ORGAMT*.

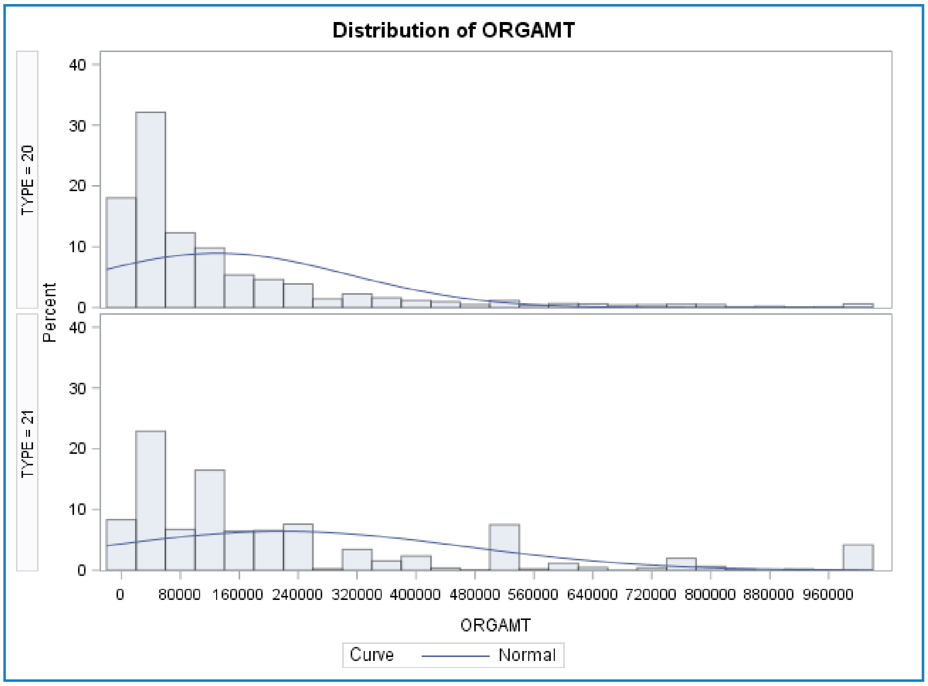

Fig.1: Distribution of ORGAMT

In examinging the data, we found the ORGAMT of 95% commercial term loans (type=20) are less than $1,000,000 and the highest ORGAMT is more than $10,000,000. Also, the ORGAMT of 90% commercial revolving line (type=21) are less than $1,000,000 with the highest ORGAMT being more than $15,000,000. The histogram shows the distribution of those loans with ORGAMT less than $1,000,000. Other loan types exibit similar patterns. With this kind of distribution, if we are to use the loan balance to calculate the default rate, a loan with a $500,000 balance that went default would give a much more significant impact to the default rate calcualtion than five loan balances of $50,000 each that went default. We call these loans with small balances being "eaten" by loans with bigger outstanding balances. With these thoughts in mind, we prefer to use the loan **Count** instead of the loan balance to calculate the default rate in our study.

Data Scrubbing and Segmentation

Current accounting standards require disclosure of loan **Segments** (ASC 326-20-50-3). There are also **Classes**, which are sub-components of Segments and generally represent different loan types requiring certain disclosures. A Class can be the lowest level of disaggregation, or **Pools** can be adopted at even more granular levels, where loans are aggregated with similar risk characteristics.

The bank already divided its loan portfolio into several segments, such as Consumer loans, Commercial loans, Real Estate loans, Construction loans, Leases, etc. These segments are further divided into different classes based on the loan types. We started by calculating simple cumulative default rates for each loan type, using the exact definition of a default as specified by the bank.

Not surprisingly, the default rates vary a lot from type to type. Some examples are shown in Figure 2: *Default Rate by Loan Types.*

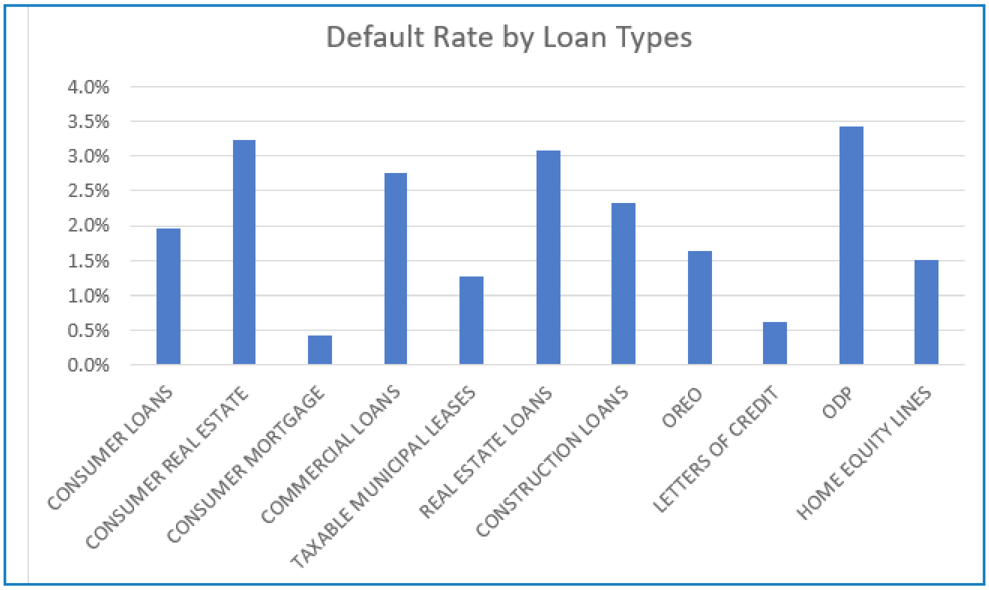

Fig.2: Default Rate by Loan Types

While the lifetime default rate of each loan type is very informative, we are also interested in the timing of default, i.e., when the default is more likely to happen. This is when survival analysis comes into play. In survival analysis, we study the survival rates/default risk along the life of the loans. As an example, we plotted the survival rates of three different types of loans: Consumer Term Loan (loan type 10), Commercial Term Loan (loan type 20) and Commercial R/E Term Loan (loan type 40), in Figure 3: *Product-Limit Survival Estimates*. We can see that Consumer Term Loan performs better than Commercial Term Loan after seven years.

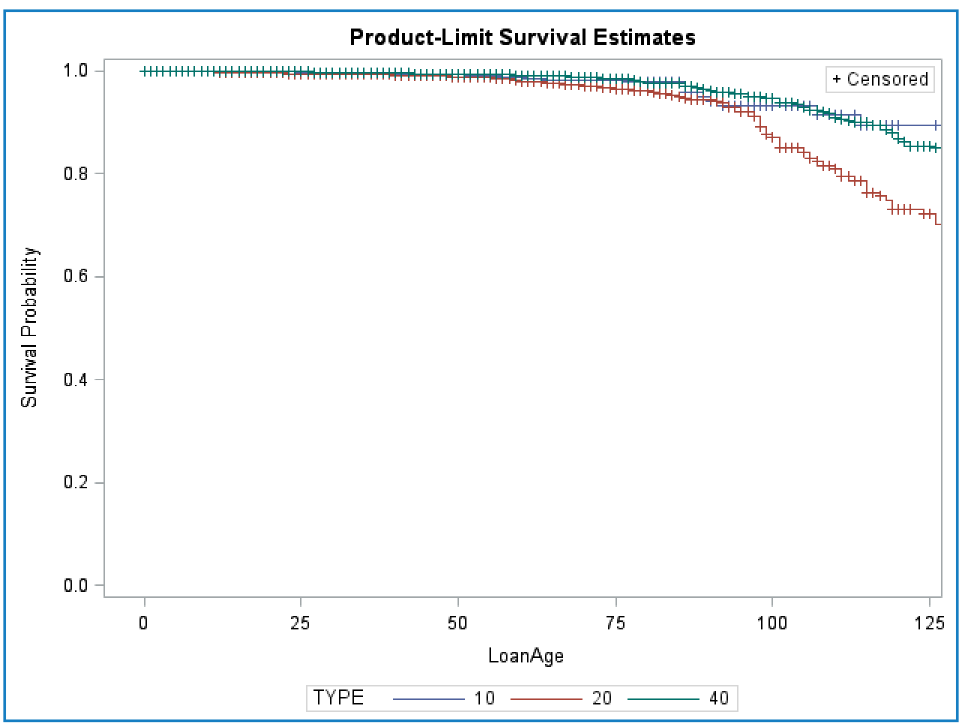

Figure 3: Product-Limit Survival Estimates

We also compared different vintages. Loans originated between 2005-2008 stand out as they performed much worse than loans that came before them. This should not be a surprise to anyone, given the U.S. economic crisis that began in 2007. See Figure 4: *Cumulative Default Rate by Vintage*.

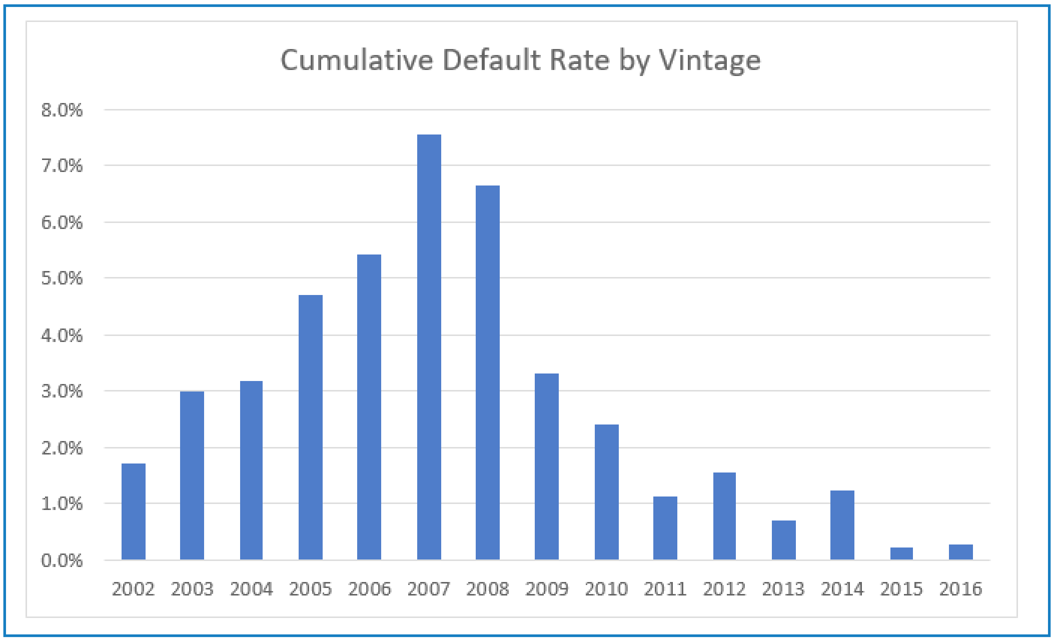

Figure 4: Cumulative Default Rate by Vintage

From 2011 through present, loan performance has been much better, with fewer than 2% of loans defaulting. Of course, part of this behavior occurs because borrowers are more likely to experience difficulties in making the payments over a longer period, so the most recent vintages will tend to have lower cumulative default rates while their loans are still relatively new. Given the average life of about 36 months of all loans, the default rate for the recent vintages is still very informative.

Default rates increased everywhere during the financial crisis, and loans with better credit ratings are supposed to fare better than others. Using the bank-assigned internal loan rating (WCHCOD), we calculate the cumulative default rate of loans in each rating group as shown in Figure 5: *Cumulative Default Rate by Original Rating*.

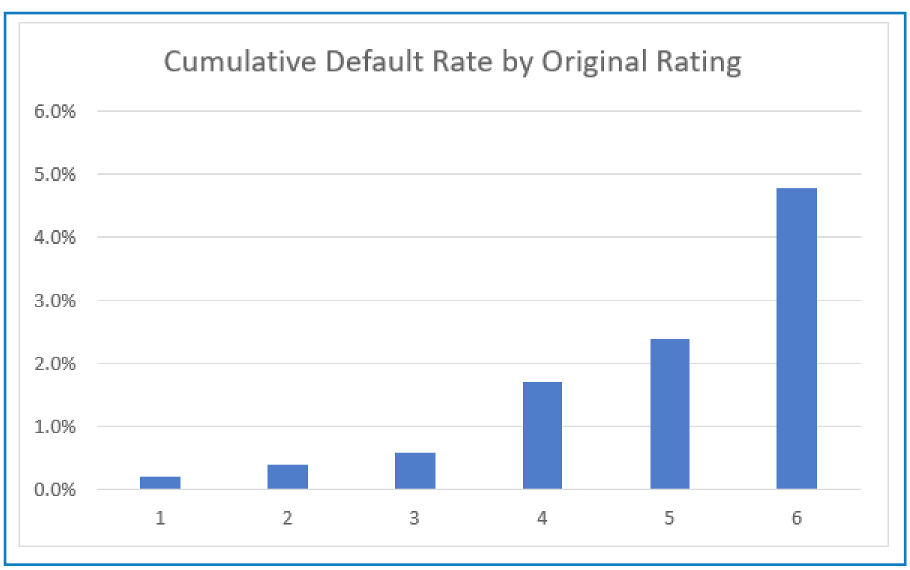

Figure 5: Cumulative Default Rate by Original Rating

Comparing Figure 4 and Figure 5, one might be curious if the ratings are normally distributed for each loan type. Table 1 shows frequency percentage of the internal ratings (dubbed WCHCOD: watch list code) of three different types of loans.

Table 1. Frequency Percentage of Internal Ratings

WCHCOD	CONSUMER TERM LOAN	COMMERCIAL TERM LOAN	COMMERCIAL R/E TERM LOAN
1	6.29	1.77	0.31
2	1.51	1.6	3.12
3	88.97	90.06	93.6
4	1.08	2.84	1.45
5	0.99	1.6	0.77
6	1.16	2.14	0.74

It turns out most loans were rated 3 among different loan types.

There are many other ways to "slice and dice" historical defaults. This analysis allows for banks to apply the loss rate method and get a better understanding of their loan performance.

HISTORICAL DEFAULT VECTORS

When there is good history data, we can also calculate the historical default vectors. These vectors can be easily applied to a discounted cash flow method (DCF) in loss analysis.

Default Rate by Loan Types

The bank's construction loans clearly are the most sensitive to the housing market condition, as seen in Figure 6: *Annual Default Rate by Loan Type*.

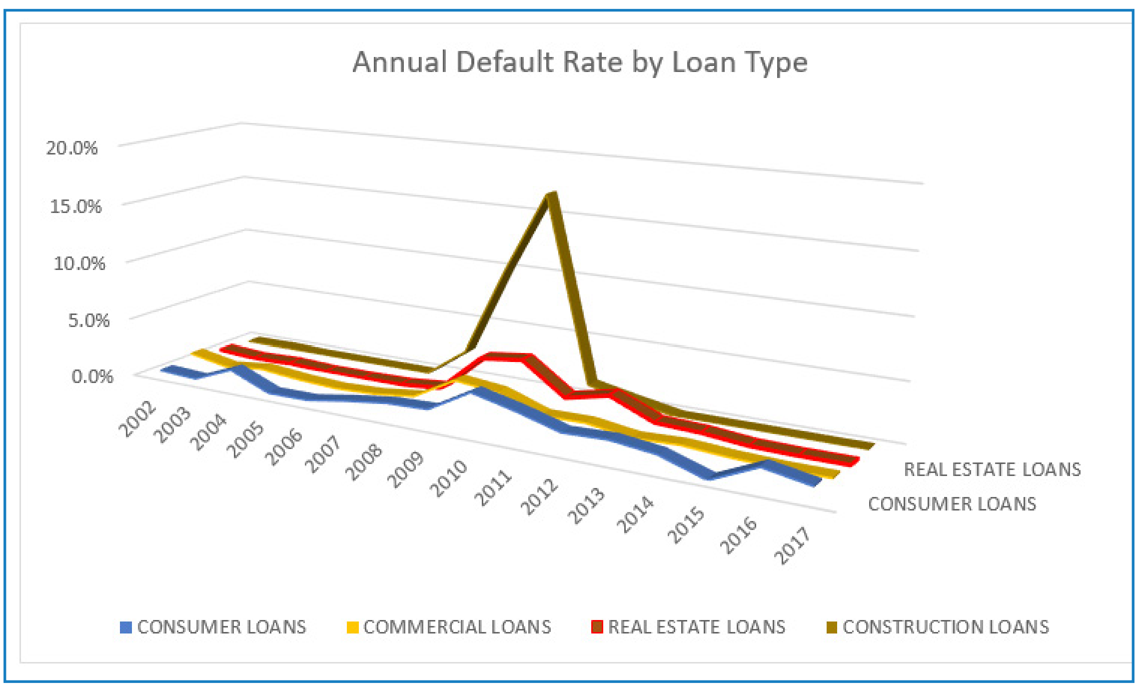

Figure 6: Annual Default Rate by Loan Type

Default Rate by Vintage

The 2007 and 2008 vintages have the highest spikes in the crisis, as seen in Figure 7: *Annual Rates by Vintage*. At the same time, the seasoned vintages such as those issued in 2002 and 2003 didn't suffer as much during that period.

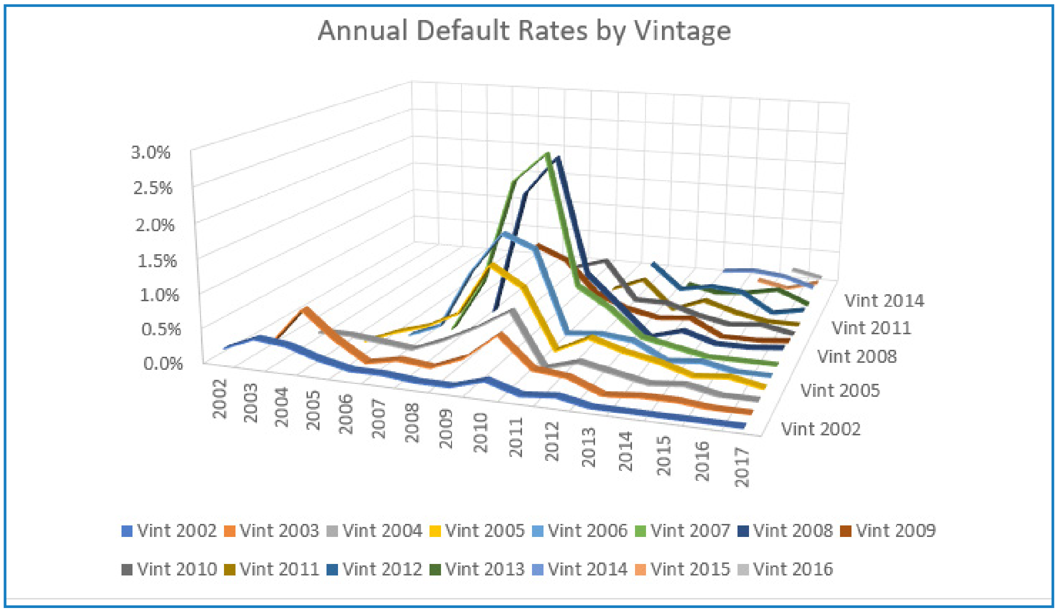

Figure 7: Annual Rates by Vintage

Default Rate by Rating

If we look at the loan defaults in each year by its WCHCOD rating, loans with a WCHCOD rating of 3 or higher clearly perform better. There is no material difference in default rates among loans in these three rating groups. See Figure 8: *Annual Default Rate by Loan Ratings*.

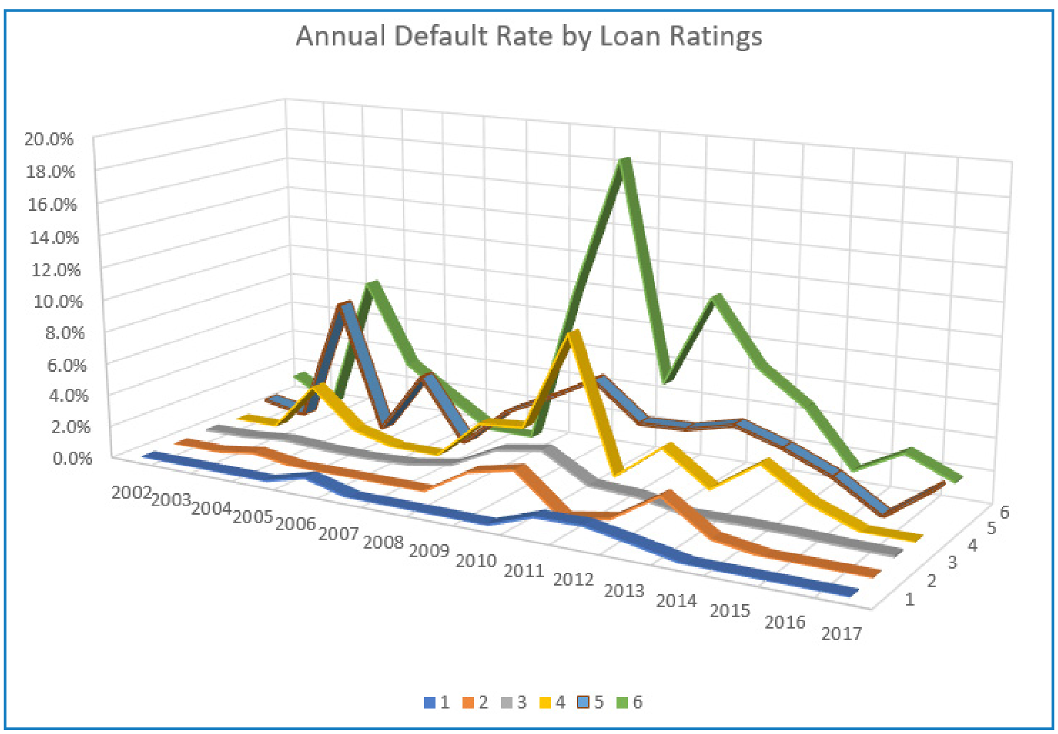

Figure 8: Annual Default Rate by Loan Ratings

Deriving these vectors sometimes requires quite some judgement; they are less complex and less data intensive than the statistical models below.

STATISTICAL MODELING

There could also be a lot of subjectivities in these models. Overall, the decisions are mostly based on statistical quantities calculated in the modeling process.

Migration Matrix

It is common in banking to estimate loan status migration probabilities. They are often displayed in migration matrices.

There is delinquency information available in the bank data set. The data are annually aggregated. We prepare the data by categorizing the loans into six different delinquency states and assume that the default event constitutes a seventh state. After some data cleaning, we developed the following migration matrix for Commercial Term Loans. See Table 2.

Table 2. Migration Matrix for Commercial Term Loans

	CURRENT	Dlq30	Dlq60	Dlq90	Dlq120	Dlq150	Dlq180	Default
Current	99.2	0.72	0.03	0.01	0	0	0	0.04
Dlq30	69.66	16.02	12.3	0.24	0.08	0	0	1.7
Dlq60	42.06	10.75	10.28	25.7	0.93	0.47	0	9.81
Dlq90	29.07	5.81	10.47	22.09	23.26	0	1.16	8.14
Dlq120	17.86	3.57	0	3.57	14.29	42.86	0	17.86
Dlq150	18.18	9.09	0	0	0	0	36.36	36.36
Dlq180	77.78	0	0	0	0	0	11.11	11.11

This matrix tabulates the estimated probabilities of migrating from a state at one month to the month following. For example, a 60-day delinquent loan whose current state is Dlq60 will have 42.06% chance to go back to the Current state in the next month, and 10.75% chance to go back to the Dlq30, and so on, as shown in the third row. As default is a terminating state, the Default state cannot go back to any other states. Hence, there is not a Default row in the table.

Also, you may notice the delinquency status change of some loans does not always make financial sense. For example, if you look at the second row of the migration matrix, 0.24% of loans that are previously 30-days delinquent are moved to 90-days delinquent after one month. This is most likely due to input error. Note that the percentage of such error is small.

The migration matrix was derived based on monthly delinquency status changes, and multiyear default probabilities can be computed by means of matrix multiplication. For example, assuming we have a loan that is 60-days delinquent, Table 3 shows its PD as well as delinquent status in the eight following periods.

Table 3. Loan PD at 60-days Delinquent Status

INITIAL STATUS		PERIOD 1	PERIOD 2	PERIOD 3	PERIOD 4	PERIOD 5	PERIOD 6	PERIOD 7	PERIOD 8
Current	**0**	42.06%	67.92%	81.07%	88.65%	94.07%	96.75%	97.93%	98.42%
Dlq30	**0**	10.75%	5.21%	2.88%	1.98%	1.39%	1.10%	0.97%	0.91%
Dlq60	**100%**	10.28%	5.64%	2.31%	1.06%	0.57%	0.34%	0.24%	0.19%
Dlq90	**0**	25.70%	9.29%	3.92%	1.64%	0.72%	0.34%	0.19%	0.12%
Dlq120	**0**	0.93%	6.89%	3.34%	1.46%	0.62%	0.27%	0.12%	0.06%
Dlq150	**0**	0.47%	0.50%	3.11%	1.49%	0.65%	0.27%	0.12%	0.05%
Dlq180	**0**	0.00%	0.52%	0.36%	1.25%	0.72%	0.33%	0.14%	0.06%
Default	**0**	9.81%	4.03%	3.02%	2.47%	1.28%	0.61%	0.30%	0.17%

The last row basically says a 60-days delinquent loan has a very high PD in the first several periods that follow, but the estimated PD reduces quickly if it survives long enough.

In ZMdesk, you can set up the migration matrix in the "Assumptions" page, see Figure 9: *Migration Matrix Set Up*, and assign it to a loan. The program will pick up this matrix at run time when forecasting the default rates for the loan.
Depending on the user's preference, migration matrices of one-year frequency or multi-year frequency can be derived in a similar fashion.

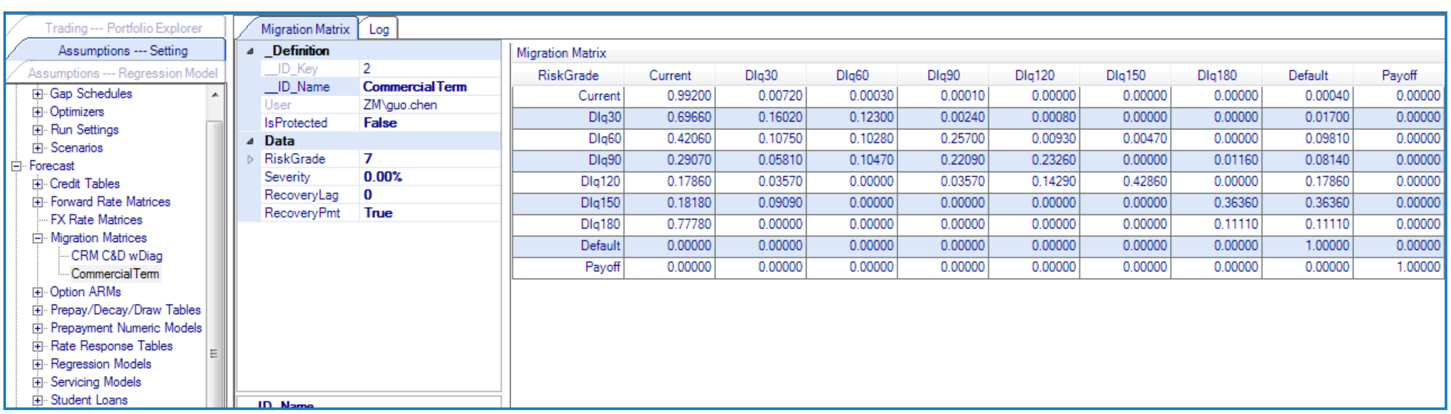

Figure 9: Migration Matrix Set Up

Logistic Regression Model

Logistic regression is a very popular regression technique due to its simplicity and good performance. Just as with linear regression, once the parameters have been estimated, the logistic regression can be evaluated in a straightforward way, contributing to its operational efficiency.

We applied the logistic regression to the Commercial R/E Term Loan portfolio. Taking borrowers, lenders, and the general economy into consideration, we started the model construction with fitting a 'full' model including all potential predictors. Only variables that can sufficiently discriminate between default and non-default are included in the "final model." It is more stable and retains the predictive power with a reduced number of variables that make significant contribution to the outcome.

The bank-assigned internal loan rating (WCHCOD) is an obvious candidate of the model. As it turns out, the first three internal ratings 1, 2, and 3 don't discriminate each other in the model. This is also supported by what we observed from Figure 5 and Figure 8. In the final model, these internal ratings are mapped to a letter-based custom rating where the internal ratings 1, 2, and 3 are mapped to the custom rating **A**; the internal rating 4 is mapped to the custom rating **B**; the internal rating 5 is mapped to the custom rating **C**; and so on. All loans with their internal ratings missing are mapped to the custom rating **M**, which stands for Missing. This custom rating serves as a much better input.

Besides the loan rating, the home price appreciation (HPA) rate is selected as the only macroeconomic input to the model. In our study, the 15 months lagged HPA yields the best model fitting result. Here, by lag we mean shifting a time series back by a certain period. In many cases, this could give better forecasts. Using the "Seasonally-Adjusted Purchase-Only Index" from the FHFA website, we compared its annual percent change (annual HPA) and the loan default events of Commercial R/E term loans. The 1- to-1.25 year lagged Seasonally-Adjusted HPA (SA HPA) series had the highest correlation with the loan defaults. Table 4 shows the correlation between loan default and some macroeconomic series under different lags:

Table 4. Loan Default Correlations

QUARTERS LAGGED	CORR. w. HPA	CORR. w. GDP	CORR. w. UNEMPLOYMENT
7	-0.651	-0.461	0.039
6	-0.721	-0.615	0.142
5	-0.769	-0.676	0.257
4	-0.766	-0.726	0.396
3	-0.739	-0.700	0.532
2	-0.686	-0.542	0.625
1	-0.601	-0.408	0.704
0	-0.540	-0.252	0.742

It's more illustrative if we put the graphs of the Commercial R/E Term Loan default events and the 15 months lagged SA HPA series side by side. There clearly is a similar pattern of spikes around 2010 and 2012. See Figure 10: *Total Default and Economic Indicators*.

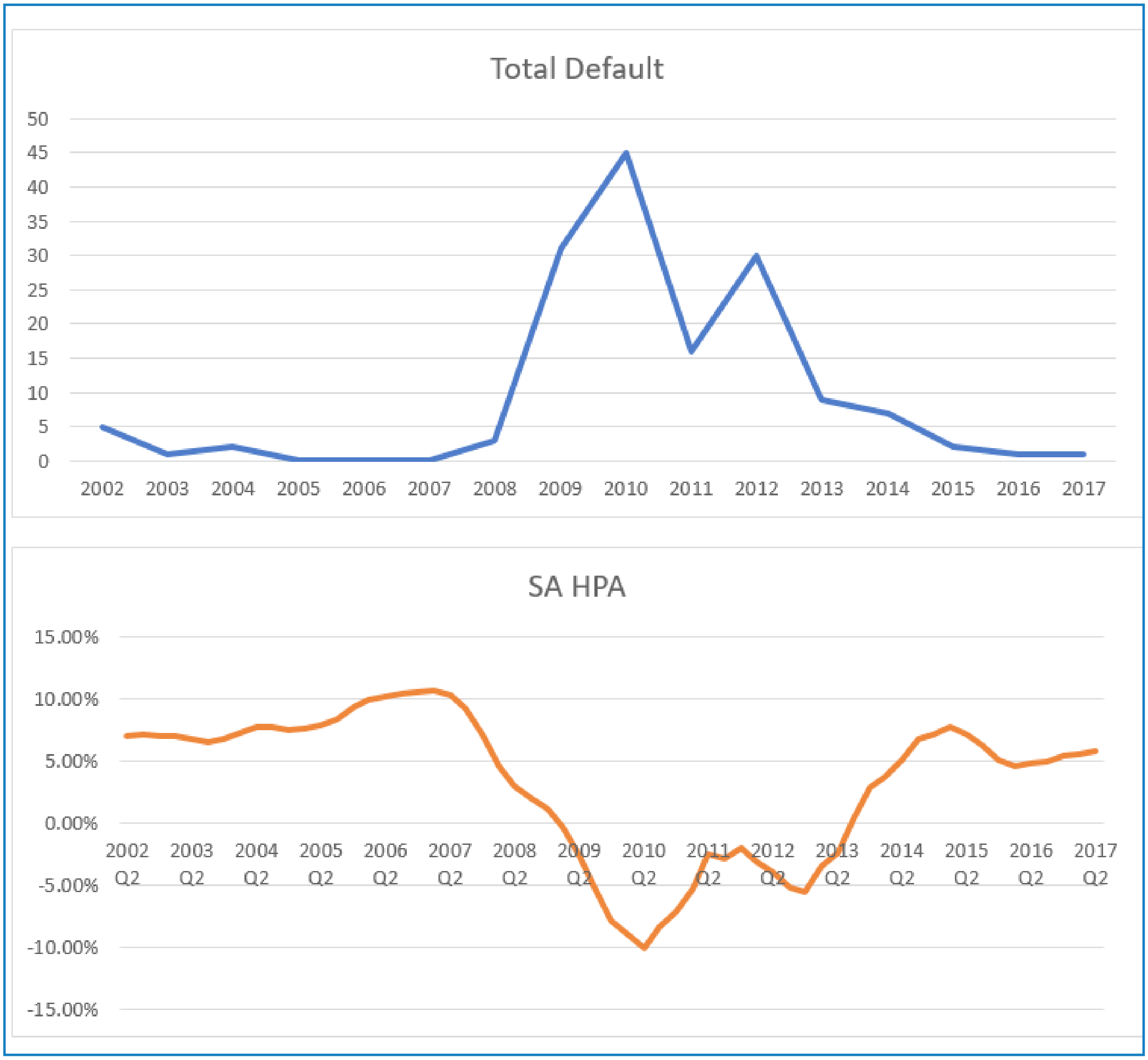

Figure 10: Total Default and Economic Indicators

To calibrate the model, the data set was first stratified into a random sample with about 70% of the Commercial R/E Term Loan data included. These in-sample loans are used for model training to estimate the model parameters. The remaining 30% out-of-sample loans are used for validation. Figure 11 provides a comparison of in-sample and out-of-sample fitting. Note PD refers to the model projected probability of default.

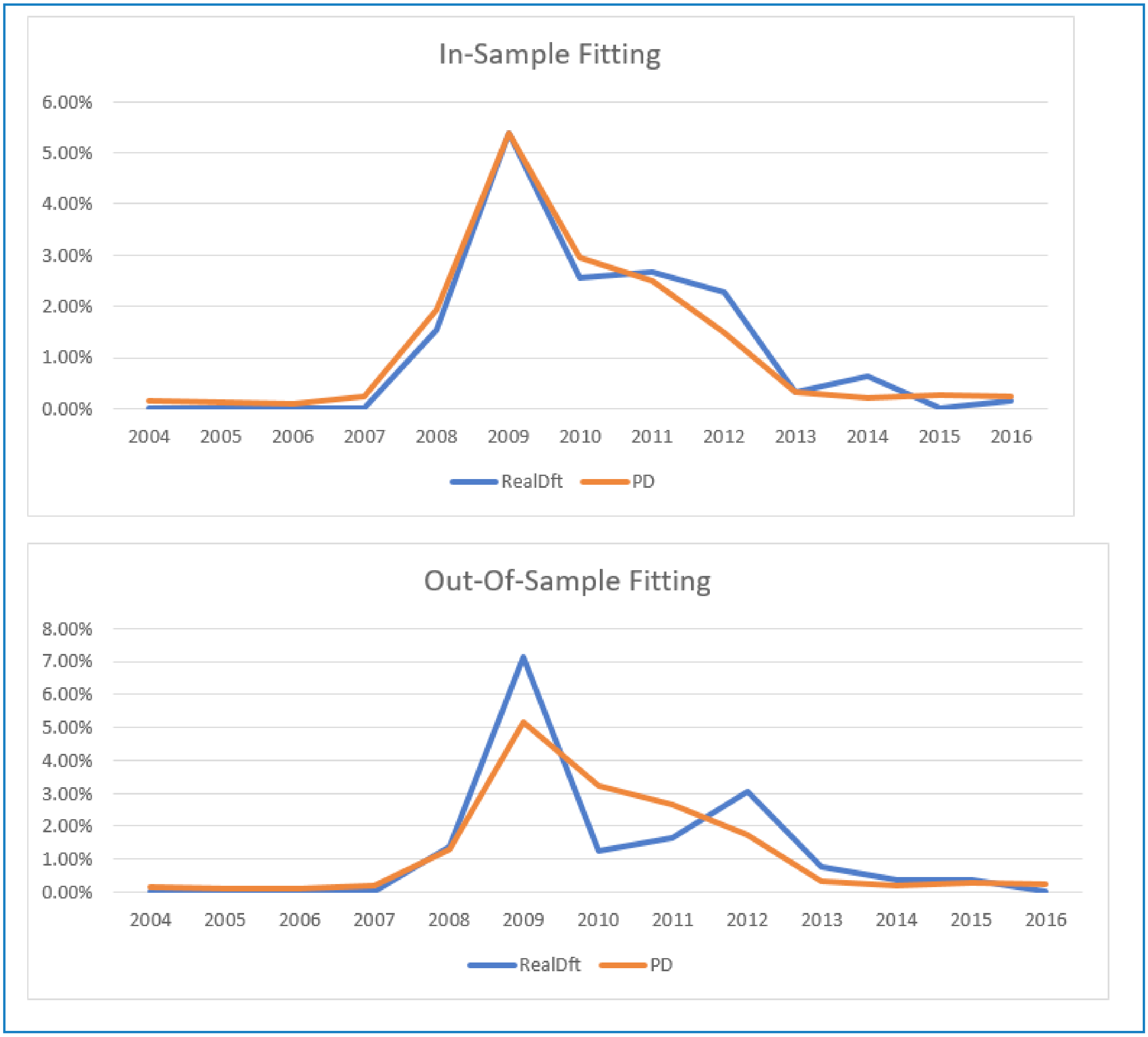

Figure 11: In-Sample and Out-of-Sample Fitting Comparison

As we said, the logistic regression can be evaluated in a straightforward way. Our logistic regression model for the Commercial R/E Loan has the following formula:

$$PD = \frac{1}{1+e^{-\left(P_0+P_1\times HPA_{Lag18}+P_2\times Rating\right)}}$$

Here $P_i s$ are the estimated model parameters; Rating indicates the corresponding $P_i s$ are the numeric estimates that are assigned to different values of these categorical variables. The model parameters are listed in Table 5.

Table 5. Model Parameters

PARAMETER	ESTIMATE	STD. ERROR	WALD CHI-SQUARE	PR>CHISQ
Intercept	-7.5914	0.1738	1907.268	<.0001
HPA_lag15	-0.2207	0.025	77.9892	<.0001
Rating B	1.7212	0.3311	27.0259	<.0001
Rating C	2.284	0.4035	32.0375	<.0001
Rating D	3.3599	0.3253	106.6648	<.0001
Rating M	6.0167	0.431	194.861	<.0001

It is not surprising that the default risk increases as HPA decreases, and better Rating reduces the default risk. Now, let us assume we have a loan issued in April 2013 with a Rating A. Using the formula above, we can manually calculate the model projected default rates as seen in Table 6.

Table 6. Model Projected Default Rate Example

LOAN SPECS		Model Parameters		Calculation Period	HPA_lag18	PD(MDR)	CDR
IssueDate	201304	**Intercept**	-7.5914	0	0.6810443	0.04%	0.52%
Rating	A	**HPA_lag15**	-0.2207	1	-0.4507042	0.06%	0.67%
		Rating	0	2	0.5592841	0.04%	0.53%
				3	1.4983352	0.04%	0.43%
				4	3.0470914	0.03%	0.31%
				5	3.0956329	0.03%	0.31%
				6	3.2560706	0.02%	0.29%

It is straight forward to set up the model in ZMdesk. Let's see how to do it from the beginning.

SET UP THE MODEL FORMULA

First, we go to the ZMdesk Regression models page to set up the formula. In Figure 12: *Model Variable Set Up*, there are already several model inputs defined in the "Independent Variables" group. They are HPA and Rating. Here, Rating is defined as the instrument property and the corresponding values will be read from the instrument at run time, whereas HPA is defined as a global time series variable whose values can be defined in the "Independent Matrices" group below.

In the "Independent Matrices" group, we can define different time series of the global "Independent Variables", e.g. HPA, GDP, etc., for different economic conditions. Here we show the HPA, Spread2Curve and Unemployment series for a "Baseline" case. Similar HPA series can be defined in the "Optimistic", "Pesimistic" cases, and you can define as many cases here as you want. Later, when we set up the run setting for a simulation, we can associate these cases with different economic scenarios. See Figure 12: *Model Formula Set Up.*

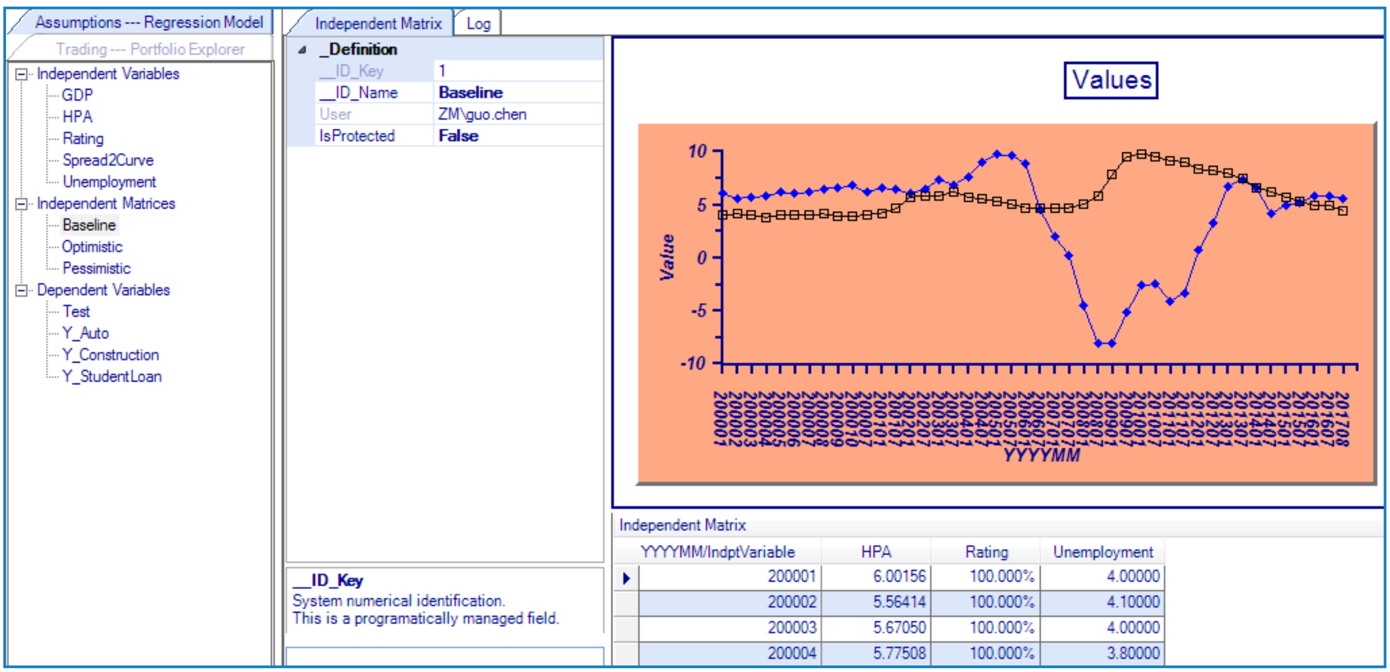

Figure 12: Model Variable Set Up

With the independent model variables ready, we can now set up the formula in the "Dependent Variables" group. Our regression model, named "Y Construction", is shown in *Figure 13: Model Formula Set Up*. This formula will be assigned to a loan later. Different loans can use different regression models in a single run.

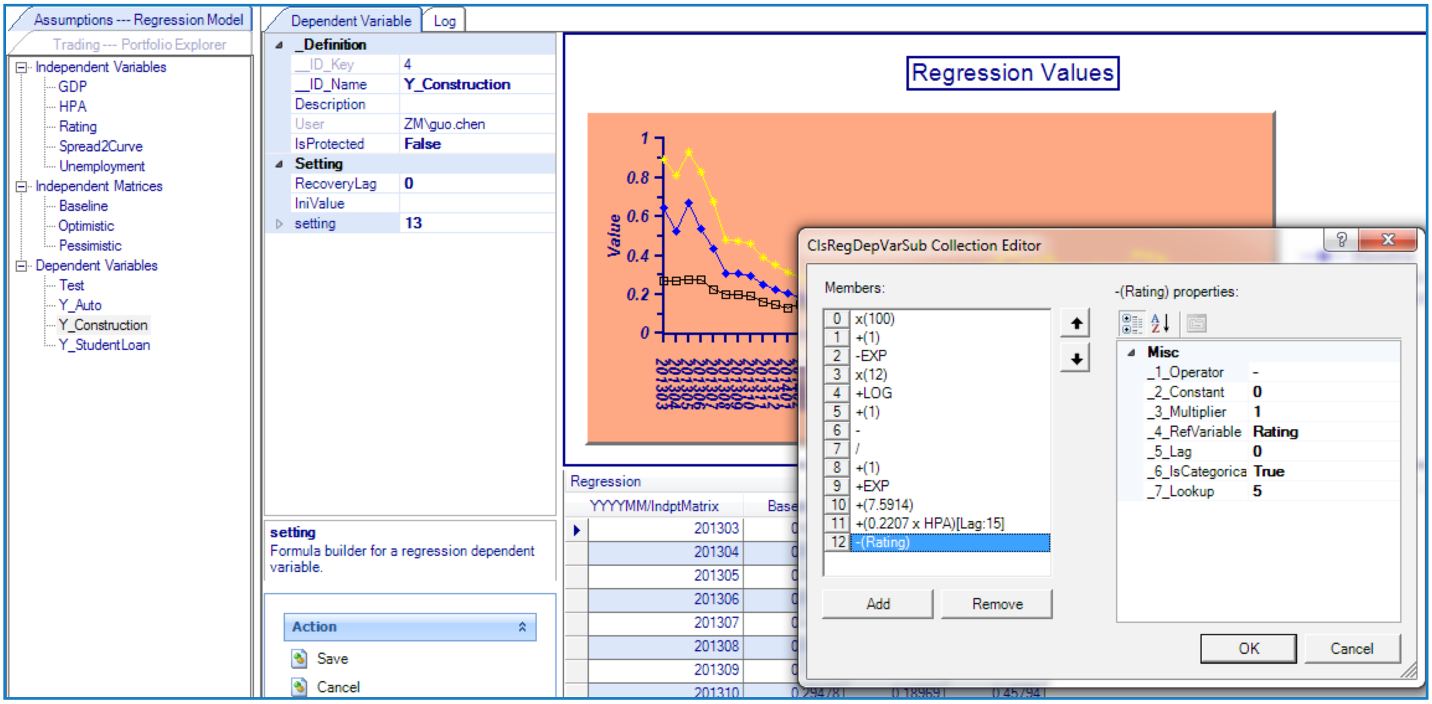

Figure 13: Model Formula Set Up

Associate the Predefined Economic Condition with the Run Setting

In the "Assumptions—Setting" page, we associate the "Baseline" economic condition defined earlier with the "_BaseOnly" scenario. See Figure 14: *Scenario Set Up*.

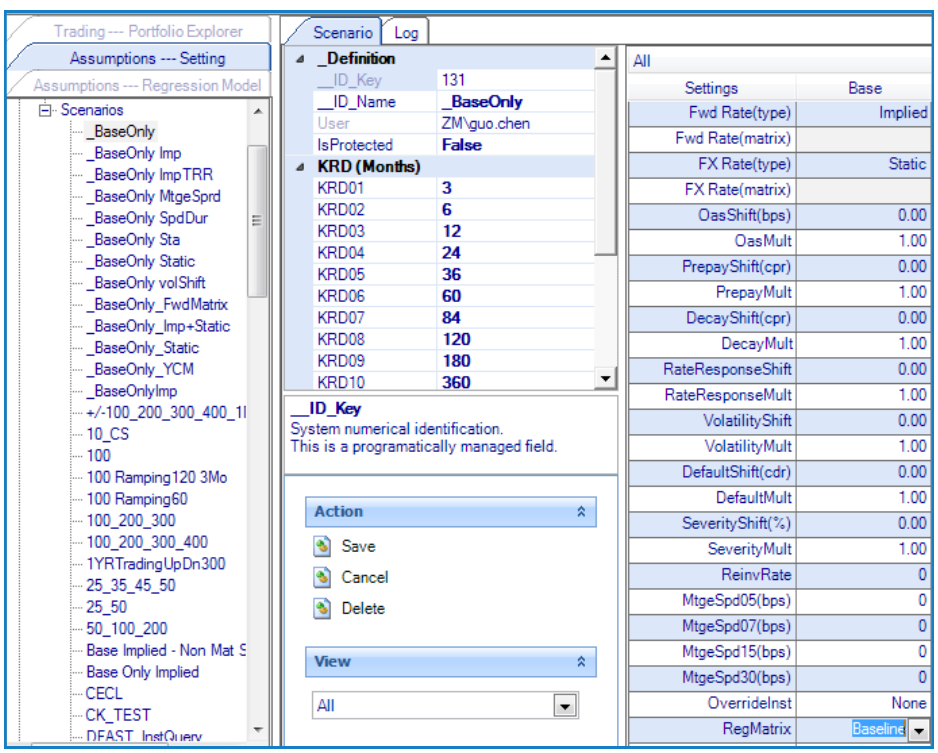

Figure 14: Scenario Set Up

Set Up the Loans

Now let's open a loan "Construction Loan1" in detail view and set its "Default" to "Regression: Y Construction", the regression formula we set up earlier. See Figure 15: *Set Up the Default Assumption for a Loan*.

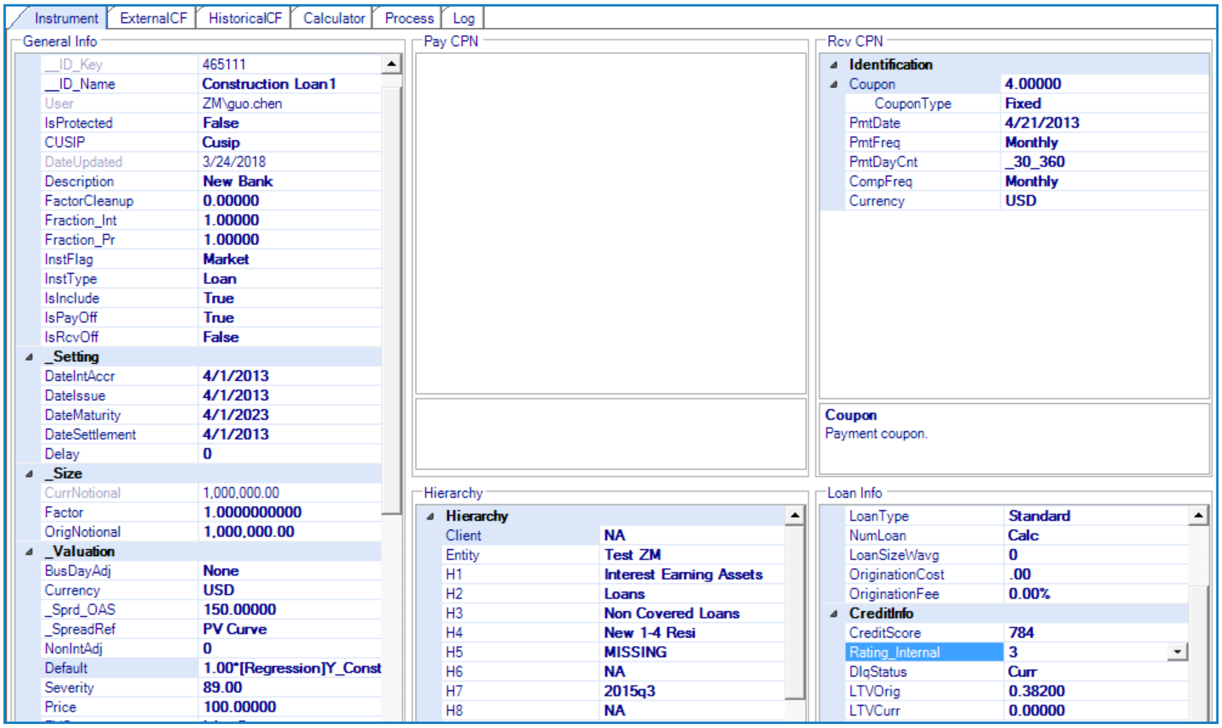

Figure 15: Set Up the Default Assumption for a Loan

Running the Loans and Result Discussion

In the "Process" page, we pick the run setting "BaseOnly" as shown below and click the "Run" button. See Figure 16: *The Process Page*.

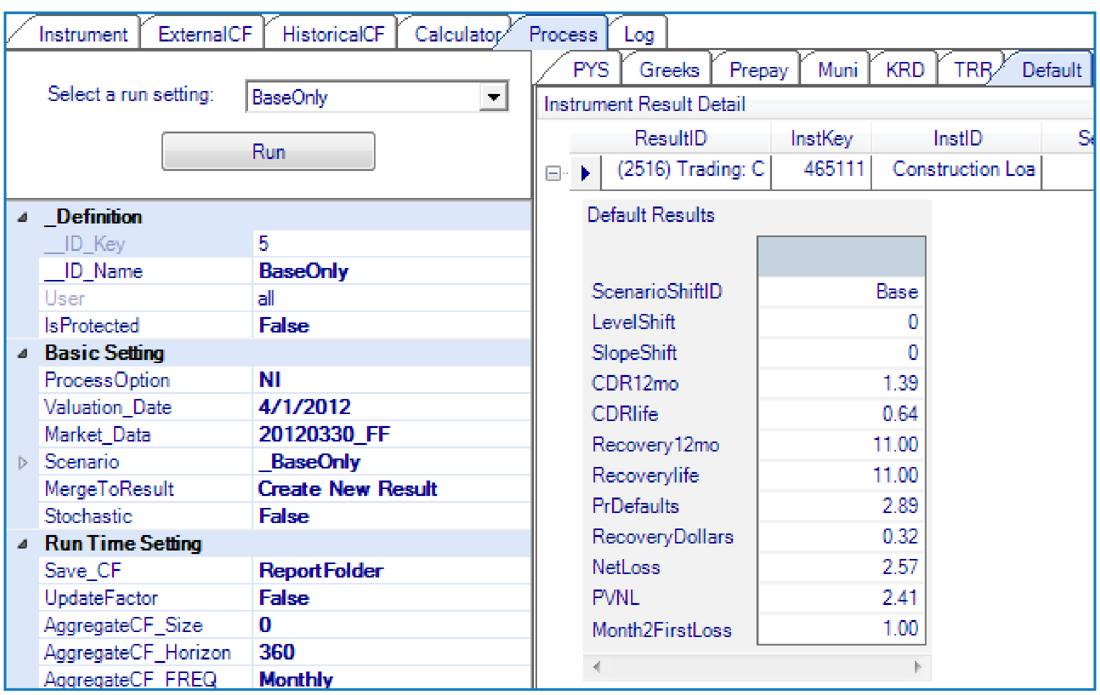

Figure 16: The Process Page

There is a lot of default related information in the "Default" tab of the Process page. Based on this information, we can get the following CECL measures, assuming a loan for $1,000,000.
CDR first 12 months – 0.34%
CDR Life – 0.2%
Total dollars in default – $8,700
Total dollars recovered – $1,000
Net loss in dollars (undiscounted) – $7,700
Net loss (discounted basis at book yield) – $7,100

You can also view the cash flow generated from the run, as shown in Figure 17. Notice the CPR and CDR columns in red. They match up with our manually calculated default rates earlier.

	A	B	C	D	E	F	G	H	I	J
1	VersionNum	ID Key	ID Name	Shift ID	CurrNotional	ValuationDate	DateIssue	PrevPmtDate	Dom	PmtDayCnt
2	2024	465111	Construction Loan1	00__Base	1000000	4/1/2013	4/1/2013	4/1/2013	21	_30_360
3										
4	**Cnt**	**Date**	**Time**	**DiscFactor**	**BeginBal**	**CPR**	**SMM**	**CDR**	**MDR**	**Severity**
5	0	04/01/2013	0.00000	1.00000000	1000000.00	0.00	0.00	0.00	0.00	0.00
6	1	04/21/2013	0.05556	0.99905947	1000000.00	5.00	0.43	0.52	0.04	89.00
7	2	05/21/2013	0.13889	0.99760150	987456.53	5.00	0.43	0.67	0.06	89.00
8	3	06/21/2013	0.22222	0.99603220	975920.03	5.00	0.43	0.53	0.04	89.00
9	4	07/21/2013	0.30556	0.99451916	964514.60	5.00	0.43	0.43	0.04	89.00
10	5	08/21/2013	0.38889	0.99295810	953209.66	5.00	0.43	0.31	0.03	89.00
11	6	09/21/2013	0.47222	0.99139839	942022.72	5.00	0.43	0.31	0.03	89.00
12	7	10/21/2013	0.55556	0.98985943	930854.40	5.00	0.43	0.29	0.02	89.00
13	8	11/21/2013	0.63889	0.98827167	919710.26	5.00	0.43	0.25	0.02	89.00
14	9	12/21/2013	0.72222	0.98673674	908615.58	5.00	0.43	0.22	0.02	89.00
15	10	01/21/2014	0.80556	0.98512857	897556.28	5.00	0.43	0.20	0.02	89.00
16	11	02/21/2014	0.88889	0.98352302	886527.72	5.00	0.43	0.18	0.02	89.00
17	12	03/21/2014	0.97222	0.98207509	875526.05	5.00	0.43	0.16	0.01	89.00
18	13	04/21/2014	1.05556	0.98045765	864557.12	5.00	0.43	0.14	0.01	89.00
19	14	05/21/2014	1.13889	0.97889492	853616.40	5.00	0.43	0.19	0.02	89.00
20	15	06/21/2014	1.22222	0.97728190	842648.48	5.00	0.43	0.19	0.02	89.00
21	16	07/21/2014	1.30556	0.97569986	831696.57	5.00	0.43	0.19	0.02	89.00
22	17	08/21/2014	1.38889	0.97406779	820760.46	5.00	0.43	0.13	0.01	89.00
23	18	09/21/2014	1.47222	0.97243628	809874.68	5.00	0.43	0.13	0.01	89.00

Figure 17: The Simulated Loan Cash Flow

Cox Proportional Hazards Model

While the logistic regression model explains the default event within a certain period, the Cox Proportional Hazards (CPH) models are regression models that link the time-to-default under consideration of censoring (meaning observation stopped due to maturity, prepayment, or default) with predictive covariates. CPH models link the default rate with a baseline hazard function and a transformation of the linear predictor (i.e., a linear combination of parameters with explanatory variables that exclude an intercept).

We investigated the Commercial Term Loan portfolio with the CPH model. Like the logistic regression model above, the custom rating acts as a good predictor. At the same time, there seems to be some quality control issue with the internal rating when a loan's original amount (ORGAMT) is relatively small. Adding ORGAMT as a categorical variable to the model improved the model fit a lot.

Besides the custom rating and ORGAMT, the 12, instead of 15 months lagged HPA also serves as a model input.

Table 7 shows the hazard ratios of different model variables.

Table 7. Hazard Ratios of Different Model Variables.

PARAMETER		ESTIMATE	STD. ERROR	CHI-SQUARE	PR>CHISQ	HAZARD RATIO
Rating	B	1.16528	0.32193	13.1019	0.0003	3.207
Rating	C	2.10422	0.26689	62.1615	<.0001	8.201
Rating	D	2.60585	0.19288	182.5195	<.0001	13.543
Rating	M	4.99388	0.30735	263.9947	<.0001	147.507
ORGAMT	Small	1.48503	0.31	22.9481	<.0001	4.415
HPA_lag12		-0.11487	0.01494	59.1408	<.0001	0.891

The hazard ratio of Rating **B** says loans with the Rating **B** have 2.2% more chance to default when compared with loans with Rating **A**. Similarly, loans with ORGAMT in the Small category is 3.415 times more likely to default. Finally, a 1% increase in the HPA rate implies a 10.9% decrease in PD.

We are interested in the performance of a typical commercial term loan under different economic conditions. Assuming there is a loan with Rating **B** and ORGAMT 65,000 (the medium original balance of the Commercial Term Loans), we set up 3 different scenarios: Downturn, Upturn, and Baseline. The Downturn scenario assumes a -3 HPA. The Upturn has a HPA of 8. And the HPA is set to 3.5 for the Baseline scenario. The survival functions of such a loan under these three scenarios are depicted in Figure 18: **Survivor Curve Under Different Economic Scenarios**.

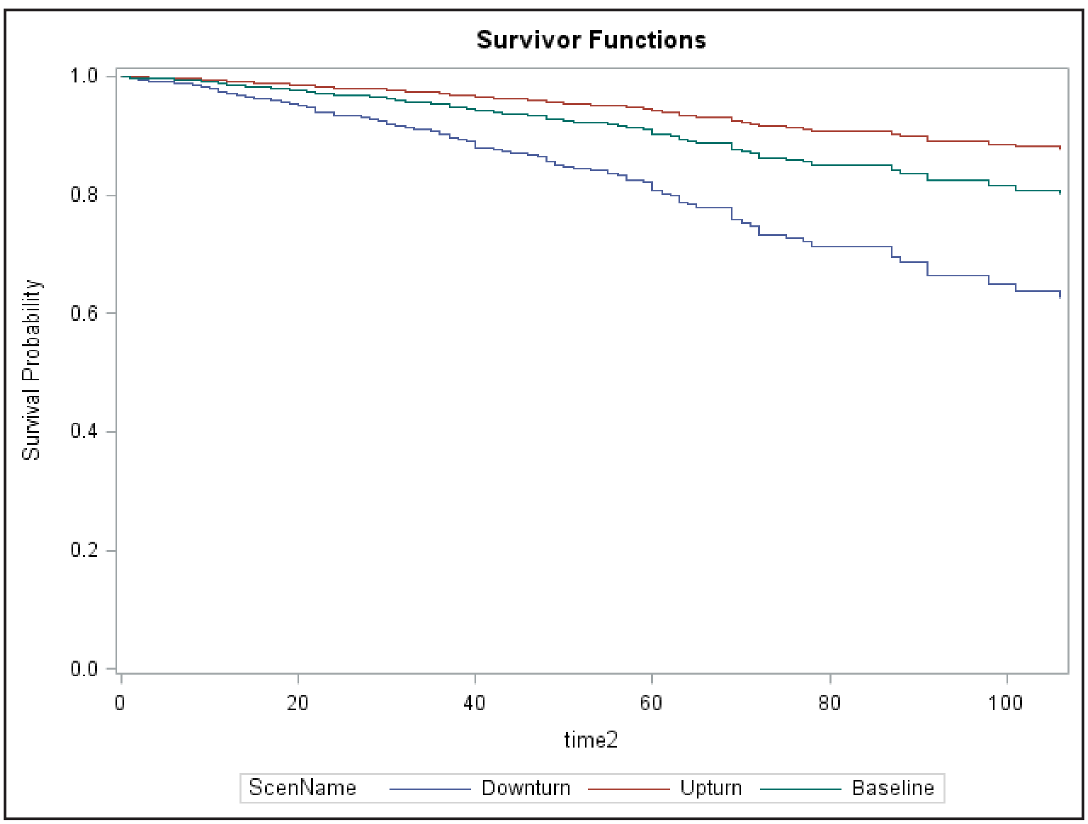

Figure 18: Survivor Curve Under Different Economic Scenarios

Figure 18 shows that, under different scenarios, the corresponding default risks all increase as time goes. More importantly, the ratios between the Survive curves give us a general idea on how different HPAs will affect the loans' survival probabilities in the model.

CONCLUSION

In this article, we showed how to analyze the bank data, choose appropriate segmentation, calculate the historical default rates, and build different statistical models that were shown to generate CECL default projections. This case study also highlights that CECL projections under the "life of the loan" concept are right at home in a robust ALM model like ZMdesk.

ABOUT THE AUTHOR

As Director, Quantitative Research, Guo Chen, PhD, is responsible for the fixed-income analytics in the solutions ZM Financial Systems provides to its clients. Chen oversees the prepayment modeling for the firm's flagship products ZMdesk and OnlineALM.com. With extensive knowledge of various interest rate term structure models, Chen leads the quantitative research team to conduct research on market trends, and perform sensitivity analysis. He is currently developing historical loan loss data modeling to help institutions respond to the upcoming CECL requirements.

Chen joined ZMFS in 2004 as a Senior Quantitative Research Associate, focusing on BondSwaps, VaR and working with 3rd party providers such as Intex, Yield Book and Bloomberg. Chen received his B.S. in Computational Mathematics from Nankai University, his MS in Optimization from Peking University, and obtained his PhD in Applied Mathematics from North Carolina State University.

