Vintage decomposition of Federal Financial Institutions Examination Council (FFIEC) charge-off mortgage data for credit risk research and education

Vadim Melnitchouk, Ph.D.
Metropoliten State University, St Paul, MN, US
Andrey Vashurin, Ph.D.
University of Telecommunications, St. Petersburg, Russia

Agenda

- Introduction
- FFIEC data decomposition
  - Naive method
  - Approximation of total losses by major vintages
- Examples of using FFIEC data
- Bridging the Academic–Practitioner Divide in Credit Risk Modeling
  - Same data
  - Same methodology

Introduction

- There is very limited number of public data sources for credit risk research and education.

- Federal Reserve database (FRED) has only two options for aggregated charge-off time series data: all banks and top 100 banks.

- FFIEC provides separate charge-off data for almost seven thousand US banks, so called CALL reports. But each quarter losses data presents an aggregation across multiple vintages.

FFIEC repository usage

- Quantitative Risk Management, Inc. (QRM) developed capital and stress testing models from CALL report data. They are using 13 Asset Types including Real Estate , Commercial & Consumer Loans.

- IMF built a model of interest rates, leverage, and bank risk taking based on variables like bank total assets, regulatory capital ratios, etc. from CALL Reports.

Why should mortgage charge-off prediction be modeling at the vintage level?

1. Vintage in mortgage credit risk is much more than just year of loan origination.

2. Vintage credit quality depends on competition, it is also a function of the supply and demand (Hughes, 2009).

Vintage data analysis in mortgage credit risk

- Mortgage default rates have spiked from historical trends in 2005 and more significantly in 2006 & 2007 beginning almost immediately after origination.

- Average time to reach maximum default rate decreased from 5-6 to 2-3 years for vintages 2006 & 2007.

Vintage level modeling

- Breeden (2011) found that consumer demand for mortgages correlated extremely well to credit quality by vintage. Demand was explained by macroeconomic drivers.

- Schelkle (2012) also built a separate mortgage default model for each vintage. His double trigger hypothesis attributed mortgage default to the joint occurrence of negative equity and a life event like unemployment.

Stochastic parametric Time to Event modeling

- Cohort (vintage) level stochastic parametric Time to Event method is well known in marketing (Hardie & Fader, 2001).

- The first application of this method in credit risk to predict the number of mortgage defaults and Time to Default distribution was presented at the OR/54 conference

  (Melnitchouk, 2012).

Vintage decomposition assumptions

- Because FFIEC repository does not have year of origination it is not possible to distinguish the separate effect of each vintage without employing some assumptions.

Method 1: Naive Decomposition

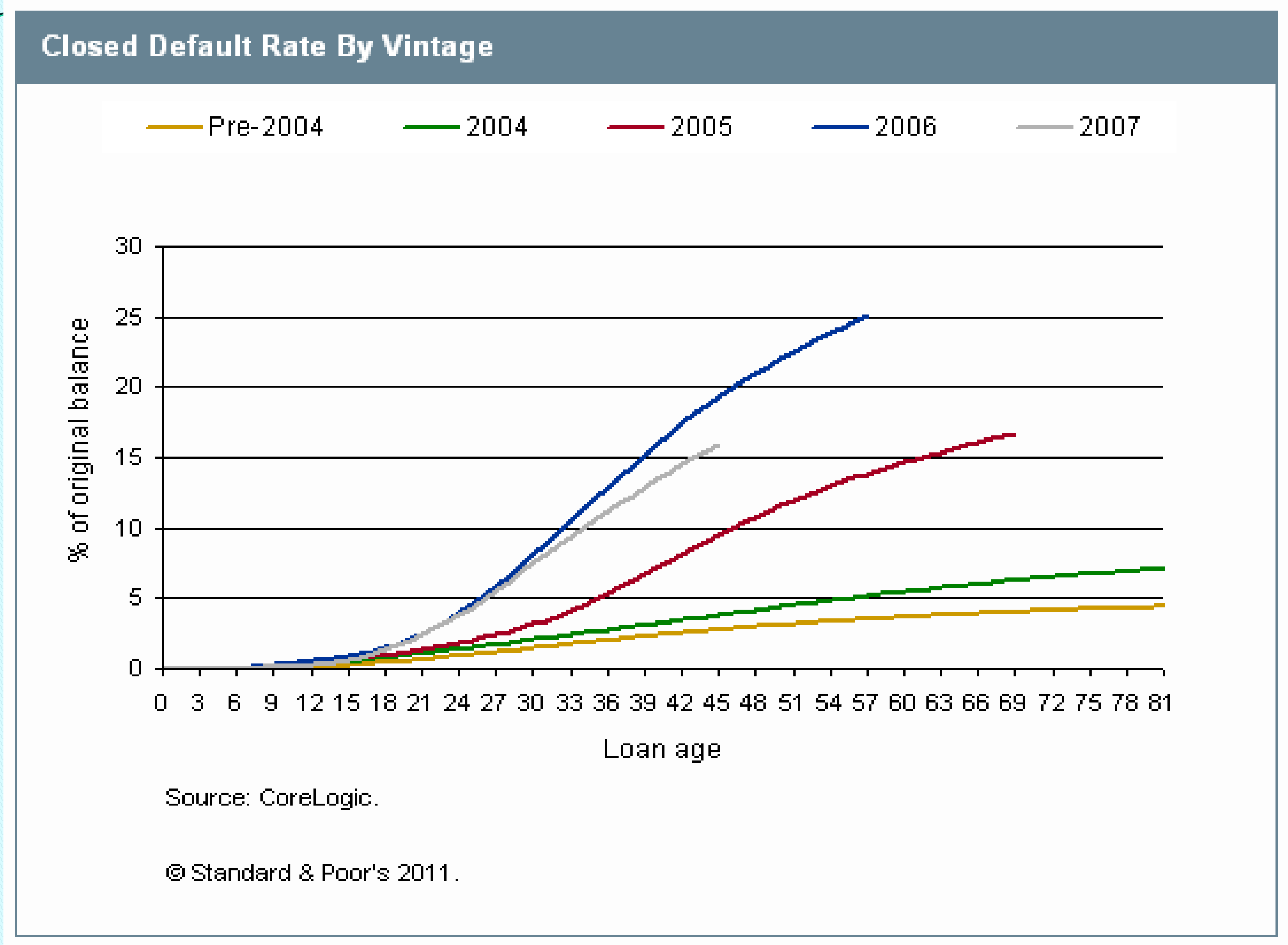
The naive method has three components:

1. US Census data on total mortgage origination amount by vintage

2. Total charge off for each bank for each quarter from FFIEC repository

3. Assumption that the same loss curve for cumulative closed default rates from Core Logic (Figure 1) is applicable for each bank.

Figure 1



Closed Default Rate By Vintage
Pre-2004
2004
2005
2006
2007
% of original balance
30
25
20
15
10
5
0
0 3 6 9 12 15 18 21 24 27 30 33 36 39 42 45 48 51 54 57 60 63 66 69 72 75 78 81
Loan age
Source: CoreLogic.
© Standard & Poor's 2011.

Naïve Decomposition steps

1. Combine global mortgage origination amount (2,773 2,908 2,726 & 2,306 billions $) for four major vintages 2004-2007 with closed default rate data (Figure 1) to get global closed default amount for each quarter for each vintage.

2. Calculate % contribution of each vintage to total default amount.

3. Apply this percentage to each bank to get default amount for each quarter for each vintage.

What is wrong with the simple method?

1. There is significant lag between closed default and mortgage charge off.

2. Vintage contribution to total charge off can vary across banks for each vintage.

Method 2: Approximation of total losses by major vintages

- To relax strong ‘naïve’ assumptions unknown ‘scale’ and ‘lag’ parameters were introduced.

- For example scale = 1.25% for vintage 2006 means 25% higher contribution of this vintage to total losses comparing to global one.

- ‘Lag’ parameter is defined as a lag between closed default (Figure 1) and charge off (FFIEC repository).

Assumptions

- According to Hermenhoff & Ohanian (2012) time spent in foreclosure can be between 5 and 15 months. Other estimations for liquidation timelines are slightly higher.

- We assume that vintage 2003 and earlier ones have significantly lower contribution to total losses during Great Recession compared with 2004-2007 vintages. Contribution of vintage 2008 (and 2009) is also ignored.

Approximation of total losses for each bank by four vintages

- Objective function was defined as minimum for the maximum difference (Chebyshev's criteria) between actual values and approximated ones.

- We estimated a total of eight scale and lag parameters.

- US Census total mortgage origination by vintage data was used to select initial values for 'scale' parameter in our optimization method.

Results

- Our research (Table 1) demonstrates significant variability in all estimated parameters. The range for 'optimal' lags is from 9 months (JPMC) to 21 ones (PNC).

- Table 2 presents # of estimated loans defaulted for WFB by vintage by quarter. We assumed average loan size of $200K.

Table 1: Total charge off approximation by vintage 2004-2007 ones

Bank	Lags, months	Scale - vintage 2004	Scale - vintage 2005	Scale - vintage 2006	Scale - vintage 2007
JPMORGAN CHASE BANK - **JPMC**	9	0.75	0.75	0.75	0.9
BANK OF AMERICA - **BoA**	12	1.11	1.25	1.25	1.25
CITIBANK	10	0.75	0.88	0.75	0.75
WELLS FARGO BANK - **WFB**	18	1.04	1.25	1.25	1.25
U.S. BANK - **USB**	10	1.25	1.25	0.75	0.76
PNC BANK - **PNC**	21	0.75	0.75	0.98	1.23

Table 2: cumulative number of defaulted loans estimation (in 1,000) by vintage for WFB

Year	Qtr	Vintage 2004	Vintage 2005	Vintage 2006	Vintage 2007
2007	2	561	547	250	18
2007	3	705	782	445	35
2007	4	914	1,225	861	57
2008	1	1,207	1,885	1,557	134
2008	2	1,346	2,370	2,242	332
2008	3	1,678	3,252	3,377	667
2008	4	2,004	4,164	4,644	1,172
2009	1	2,373	5,161	6,158	1,899
2009	2	2,823	6,373	8,036	2,944
2009	3	3,293	7,541	9,894	4,045
2009	4	4,054	9,492	12,733	5,725
2010	1	5,255	12,403	17,039	8,254
2010	2	6,059	14,663	20,383	10,496

Using FFIEC data. Example 1: Stochastic parametric model

- According to Bellotti & Crook (2007) survival (hazard) modeling is competitive alternative to logistic regression when predicting default events.

- The method has become a model of choice in recent publications.

- We apply, well known in marketing (Fader & Hardie, 2007) stochastic parametric ***Time to Event*** method to build a simple model.

Assumptions & inputs

1. Weibull distribution for Time to Default

2. Gamma distribution default density across obligors to include unobserved consumer heterogeneity.

3. Vintage level modeling to avoid aggregation bias

Inputs: Monthly # of defaults in 2006-2009 and Time varying covariates: unemployment and Home Price Index (HPI).

Model training and validation

1. Model training period for vintage 2006 was June 2006 – March 2009.

2. April 2009 to March 2010 period was selected for 'out of time' validation because unemployment increased from 8.5% to 10.1% during this period.

3. The model was implemented in MS Excel (using Solver) and in SAS/IML. Maximum likelihood was estimated to get values for five parameters.

Data

1. US Census total amount mortgage origination in 2006

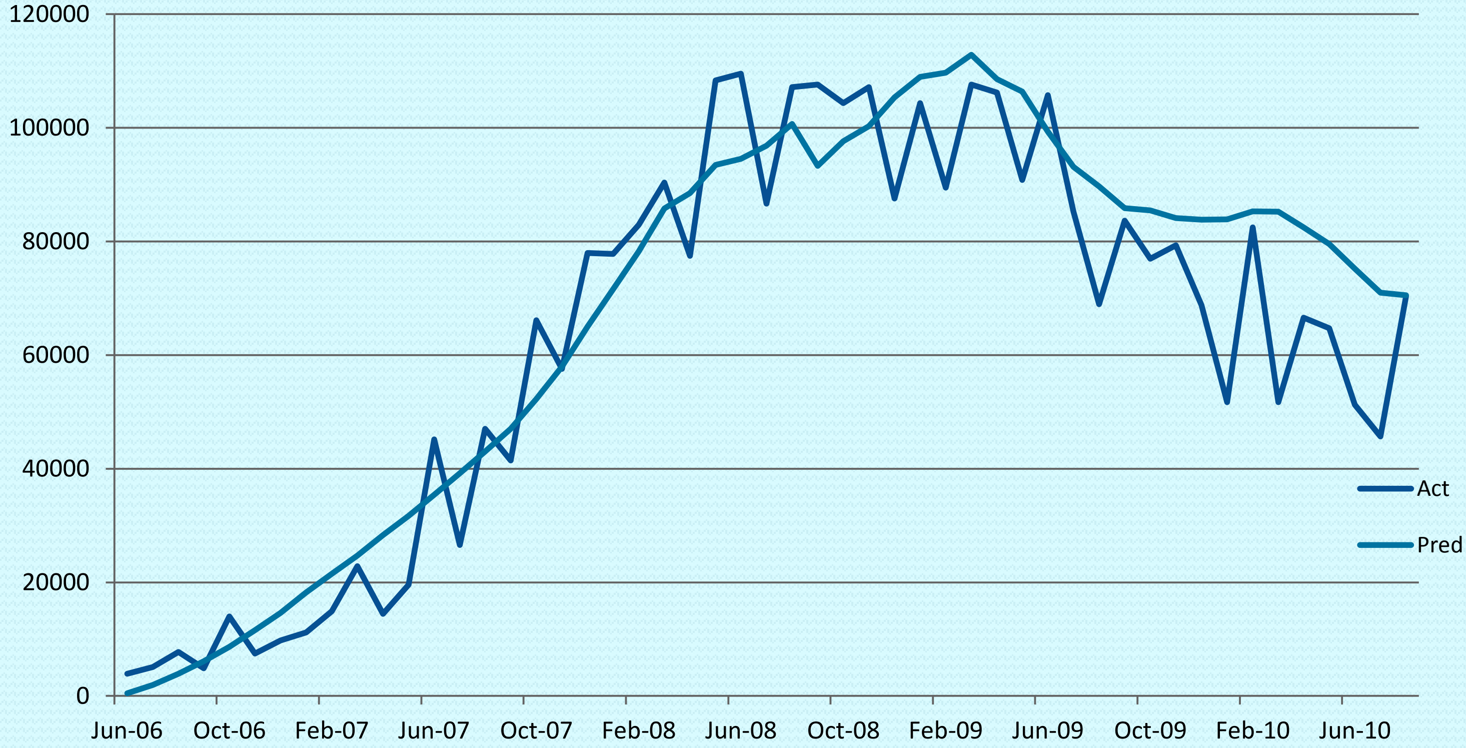
2. CoreLogic closed default rate for 2006 vintage

3. Average mortgage loan $200K assumption

Actual and forecasted values are presented at Figure 2.

Figure 2: Forecasted vs Actual monthly # of defaults

Weibull/Gamma model for 2006 mortgage origination year (All banks).



Results: Parameter Estimates

- Out-of-time forecast accuracy was at acceptable level. It is conservative enough to satisfy regulators.

- Following Crook , Bellotti & Leow (2012) we built a model without frailty assumption by omitting Gamma distribution.

- By moving to Weibull only model with four parameters we failed to find good in-sample fit.

Example 2: simple two states Markov Chain model using FFIEC data

- We assumed just two states: current loans and all stages of delinquency including charge off.
- Such simplistic model (with two unknown parameters) does not have adsorbing state.
- Steady state distribution between current and delinquent states was calculated using 2010-2012 data (Table 3).

Table 3. Delinquency steady state distribution for two states first order Markov Chain model

Bank	Steady state distribution, Current %	Steady state distribution, Delinquent %	Time to reach within 10% of steady state, in quarters
All banks	89.64	10.36	19.9
BoA	76.97	23.03	5.9
WFB	81.86	18.14	13.6
USB	90.17	9.83	11.3

Results

- We found a range of mortgage portfolio quality (measured as percentage of delinquent loans at steady state) from 9.83% for USB to 23.03% for BoA.

- Time to reach within 10% of steady state is the highest for WFB (13.6 quarters) and the lowest for BoA (5.9 quarters).

- It will take 19.9 quarters for a dynamic system which includes all banks to reach steady state.

Sensitivity to model's structure

- Because FFIEC repository has both, delinquent and severe delinquent loans, a slightly more complicated model can be developed.

- We can compare results of two models for the same bank to analyze model's robustness.

Table 4. Delinquency steady state distribution for three states for first order Markov Chain model

Bank	Current	Delinquent	Severe Delinquent
All	91.52%	2.35%	6.13%
BoA	74.54%	3.16%	22.30%
WFB	84.02%	3.74%	12.24%
USB	94.39%	1.20%	4.41%

Results

1. Lower than average sensitivity (to model complexity) for BoA

2. Slightly higher than average sensitivity for WFB

3. Much higher than average sensitivity in case of USB

Bridging the Academic-Practitioner Divide in Credit Risk Modeling

- This is a simple solution to use the same data and the same methodology.
- It requires optimal complexity model.
- Loan level models used currently by practitioners for mortgage default forecast with unemployment are struggling with aggregation bias.

Aggregation bias

- Aggregation bias is, generally, the incorrect assumption that "what is true about the group is true about the individual" (Wikipedia).

- Usually banks do not have data regarding Customer current employment status.

- It means that any loan level model (survival or regression) will assign two Customers with identical internal variables like FICO but different employment status extremely high Probability of Default during Great Recession.

Optimal complexity model

- Vintage level model which predicts # of defaults (Fader & Hardie, 2007) does not have such bias.
- We recommend practitioners to avoid aggregation bias by aggregating loan level data and building vintage level model.
- At the same time, by using vintage decomposition of FFIEC data academics can build loss forecast model at the same vintage level as practitioners.

How academics and practitioners can benefit from FFIEC data decomposition by vintage?

- Academics and Practitioners can use the same datasets to build vintage level loss forecast models.

- A FFIEC data based model can also be used for benchmarking 'internal' models developed by financial institutions.

Benchmarking stress testing models

- Section 165 of the Dodd-Frank Wall Street Reform and Consumer Protection Act directs the banking regulators to order banks with assets over $10bn to conduct stress testing under various scenarios.

- As a fast solution vintage decomposition of FFIEC data can be used for banks in this category (like First Niagara Bank, TCF bank, etc.) to build accurate loss forecast models.

Conclusions

1. Vintage level predictions for FFIEC mortgage data can be used in research and education

2. Loss forecast by vintage modeling can also be applied by practitioners for benchmarking their Stress Testing Models

3. There is a chance that by using the same data and the same level of aggregation academics and practitioner can reduce current gap in Credit Risk Modeling

Future Lines of Research

1. Vintage decomposition of FFIEC mortgage delinquency data.

2. Building Markov Chain for each vintage for major banks to estimate steady state distribution.

3. Inclusion of macroeconomic variables in non stationary Markov Chain models for benchmarking CCAR models.