The Deep Future Analytics CECL Study: Alternatives, Impacts, Accuracy, and Complexity

Joseph L. Breeden, Principal Investigator Updated April 4, 2017

This study was sponsored by Allied Solutions, NAFCU – National Association of Federal Credit Unions, and OnApproach. Any opinions expressed here are solely those of the author and may not represent the opinions of the sponsors.

Executive Summary

The new accounting rules for estimating loan loss reserves offer general guidelines and a list of possibilities, but no specific recommendations for how best to implement those rules. The present study uses a large mortgage dataset from Fannie Mae and Freddie Mac to test a range of models and options. The results quantify the pros and cons of these options.

Study Design

For the underlying models, the following were tested: time series correlations to macroeconomic data, roll rate models, vintage models, state transition models, and discrete time survival models. These models were assessed for accuracy, robustness to small data sizes, complexity, computation time, and procyclicality of lifetime loss estimates.

In all cases, scenarios were created with 24-month macroeconomic history followed by mean-reversion to long-run macroeconomic conditions. Undoubtedly, many practitioners will create two separate models, a near-term model with a macroeconomic scenario and a long-run through-the-cycle loss model. Using a single model with a mean-reverting macroeconomic scenario is preferable, because the active portfolio is used for the lifetime loss forecast rather than an average of past portfolios. It also avoids the need to validate two separate models.

The guidelines also mention the option of using a discounted cash flow approach. DCF is not a model so much as a system of equations for aggregation, since it requires estimates of default and attrition probabilities as estimated in the models tested here. Therefore, all model results were shown as direct loss aggregation, discounted loss aggregation, and DCF aggregation of cash flows simulated from the loss estimation models.

Results

The following results are intended to be used to assess trade-offs in CECL implementation details.

Foreseeable Future

Using mean-reverting scenarios here allowed the model to adapt to the current portfolio for the lifetime estimation rather than use an average over past portfolios, but at greater complexity. Conversely, it requires only one model rather than two. Even though most practitioners will use a through-the-cycle average default rate as the long-run model, we know from Basel II that these are actually models with their own complexities in estimation.

Accuracy

Projecting losses via time series models of default and pay-down rates produced an average 3-year cumulative error rate of 17-19%. In itself, that will raise concerns with validators, but the accuracy is unchanging relative to the amount of training data, which can be useful for very small or noisy data sets. Vintage models were consistently high performers in terms of accuracy with 1% to 3% error rates. Discrete time survival models and state transition models both perform well (3.2% to 6.5%), but not better than vintage models, showing that loan-level modeling does not guarantee more accuracy. Vintage, state transition, and survival models all had similar scaling properties versus size of training data. Roll rate models were consistently the worst performers at 15% to 20% error rates. Moving averages of historic loss rates are unsuited to lifetime loss forecasting at 60+% error rates. Overall, roll rate and historic average models should not be used for long-lived products.

Creating separate models by US state did not provide greater accuracy when compared to a single national model of the same portfolio. Geographic segmentation provides advantages in business application but not model accuracy.

The guidelines state that vintage modeling is not a requirement. If we assume that "vintage model" refers to any approach that adjusts credit risk and prepayment risk based upon the age of the loan, then the results show significant increases in accuracy for techniques incorporating this (vintage models, state-transition models, and discrete time survival models) as compared to those that do not include it (time series and roll rates).

Accuracy vs. Complexity

The loan-level models (state transition and survival) were by far the most complex in terms of numbers of coefficients and computational time. This complexity did not provide any increased accuracy relative to vintage models, but it does provide business value in account management, collections, pricing, and strategic planning.

The added complexity of roll rate models when compared to time series models provided little benefit other than the change to be more accurate for the first six months of the forecast. Vintage models were the overall winners in the accuracy versus complexity trade-off, so long as sufficient data exists for robust estimation.

Optional DCF

Starting from a lifetime loss forecast, using a time-value of money discounting of the projected monthly losses using the par rate on the mortgage results in a 20% to 30% decrease in the reserve amount. Estimating the principle and interest payments adjusted for the risk of default or prepayment from the loss model and then discounting with the par rate on the mortgage results in an equivalent reduction in the loss reserve as compared to the original lifetime loss forecast.

Old vs. New Rules

The magnitude of the change from the old loan loss rules to CECL will depend strongly on the lifetime of the asset and the point in the economic cycle when the adoption occurs. For 30-year fixed mortgage, the average life of loan is about 5.5 years and the lifetime loss reserve will be 4 times a historic average approach with 24 month loss emergence period. If adoption had occurred just before the onset of the last recession, the adjustment would have been 10x. At the peak of the recession the change would have been 2x. Well into recovery they would have been at parity.

Conclusion

By design, the new CECL rules provide a significant amount of flexibility in implementation. As seen from this study, even with a straightforward product like 30-year fixed rate conforming mortgages, the range of models listed in the CECL guidelines can produce a range of lifetime loss numbers that vary by a factor of 2. With the option of discounted cash flows, then the range of final answers would vary by more than a factor of 2.

Being able to choose options that will create such different answers will put the burden on lenders not only to choose the most appropriate models for their portfolios, but in doing so to also choose the level of loss via the models chosen, and to defend that choice to validators, auditors, and examiners.

The Deep Future Analytics CECL Study: Alternatives, Impacts, Accuracy, and Complexity

Joseph L. Breeden, Principal Investigator March 16, 2017

This study was sponsored by Allied Solutions, NAFCU – National Association of Federal Credit Unions, and OnApproach. Any opinions expressed here are solely those of the author and may not represent the opinions of the sponsors.

Introduction

The release of the new rules for loan loss reserves accounting by the Financial Accounting Standards Board (FASB) will bring many changes to the lending industry and no small amount of uncertainty about how to comply with the new guidelines. The Current Expected Credit Loss (CECL) framework [FASB 2016] seeks to apply forward-looking methods to estimate credit losses for setting loss reserves. This represents a dramatic departure from traditional methods of estimating the Allowance for Loan and Lease Losses (ALLL) that relied upon averages of historic losses.

These changes were obviously necessary following the US Mortgage Crisis of 2009. Figure 1 compares loss reserves to the forward-looking 12 months of losses.[1] If under the previous rules loss reserves were intended to cover the next 12 months of losses[2], clearly this did not happen. Loss reserves barely registered the 2001 recession and did not peak until a year after the 2009 recession. In all cases, reserves can be seen to be too low entering a recession and too high afterward, because the reserves were backward-looking to the previous phase of the economy. Reserves estimated under CECL will still not be perfect, because no one has a perfect view of the future of the economy, but the loss reserve estimates can be much better than before.

[1] FDIC data on the banking industry is available at www.fdic.gov/bank/analytical/qbp.

[2] A loss emergence period of 24 months is more appropriate for mortgage, as discussed later, but credit unions typically use 12 month periods for loss reserves.

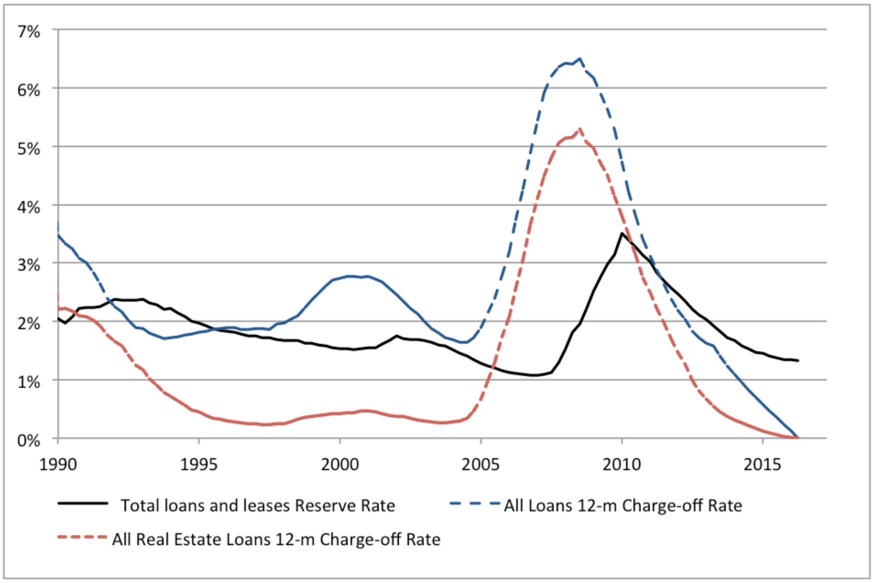


Figure 1: A comparison of industry loan and lease loss reserves to forward-looking 12-month cumulative actual losses for all loans and just real estate loans. Data from FDIC.

The final guidance from FASB indicates that financial institutions may use different estimation methods for CECL depending upon the size and sophistication of the institution. In fact, this is really a product-by-product decision, because an institution will have dramatically different volumes by product. The flexibility under CECL brings benefits to the institutions, but it will also bring a significant amount of confusion. Under the published guidance, all institutions will need new models to satisfy the rules, even if those models are simplistic for the smaller institutions. Institutions will need to determine for themselves how much sophistication they need or want. An even greater problem may come from examiners, auditors, and validators trying to review these models. They will also be trying to gauge how much sophistication the lender should have employed.

The primary goal of the current study is to create a public document that all parties can use as a reference to the strengths and weakness of the available techniques. This assessment was not a search for the "best" answer. Rather, we assessed the range of answers produced by different techniques in the context of accuracy, complexity, volatility, and stability for portfolios of different sizes. Of course, modelers, validators, and examiners will continue to debate what is appropriate, but practitioners may leverage this study as a data point on an important asset class.

Project Design

This study analyzed large datasets from Fannie Mae and Freddie Mac on conforming mortgage performance. The same data was modeled with the most commonly utilized loss forecasting techniques that can be made to satisfy the final CECL rules.

The study included a comparison of the level and timing of the predicted loss reserves through the last recession, model accuracy tests, volatility of the estimates through time, stability of the techniques as the size of the data set is reduced, and an assessment of the implementation complexity.

The greatest advantage of this study is the consistency of model creation across the various techniques tested. The same data set spanning the same time frame with the same segmentation was employed throughout. The same economic scenarios were used for all models.

Current Expected Credit Losses (CECL)

CECL was developed concurrently with IFRS9, the new accounting standard from the International Accounting Standards Board (IASB) also with the intent of creating forward-looking loss reserves. FASB elected to use lifetime loss estimation for all loans rather than having a multi-stage approach as present in IFRS9. This decision was to simplify the process for the thousands of smaller lenders in the US. The stated goals of lifetime loss estimation, using current economic conditions for the near-term, and relaxing onto the long-run economic conditions for the rest of the forecast naturally lead to the same kind of models discussed above for IFRS9.

In an attempt to soften the burden for smaller institutions, the CECL guidelines state explicitly that complex models are not required for smaller institutions. Nevertheless, finding a simpler approach that does not carry a harsh penalty in loss reserve levels or in auditor and examiner review is not obvious.

The CECL guidelines provide the following principles that should be considered when estimating losses.

Lifetime Loss Estimation

Probably the biggest change in CECL is the adoption of loss reserves for the full lifetime of the loans. Under the previous FAS 5 rules, loss reserves were computed only for those losses that are currently experiencing a loss-triggering event, whether observable by the lender or not. The loss estimation would cover the period of time required for those distressed accounts to charge-off. Typical examples are job loss, bankruptcy, large debt growth, etc. That time span is referred to as the Loss Emergence Period (LEP), Loss Discovery Period (LDP), or Incurred Loss Period (ILP). The requirement to determine a product-specific LEP was a refinement over the previous standard of only holding loss reserves for the next 12 months.

However, being defined as an event unobserved by the lender makes the LEP very difficult to estimate, usually falling to judgment and rough estimates.

McPhail, et. al., [2015] offers the following guidance,

> "… mortgage products typically have an LEP of approximately 21 to 24 months, while other consumer products usually have an LEP of 12 to 18

months. Commercial LEPs can vary considerably, depending on the product types and workout periods, but are usually greater than 12 months. Generally, institutions should use at least 12 months for their LEP. "

With this in mind, when we compare CECL to FAS 5 results, we will use a 24-month LEP for FAS 5 estimates.

"Lifetime loss" is measured from the present age of the loan to its furthest non-cancellable end point, Figure 2. For fixed term loans, this is easily determined. Lines-of-credit or renewable loans are more complex but are not included in this study.

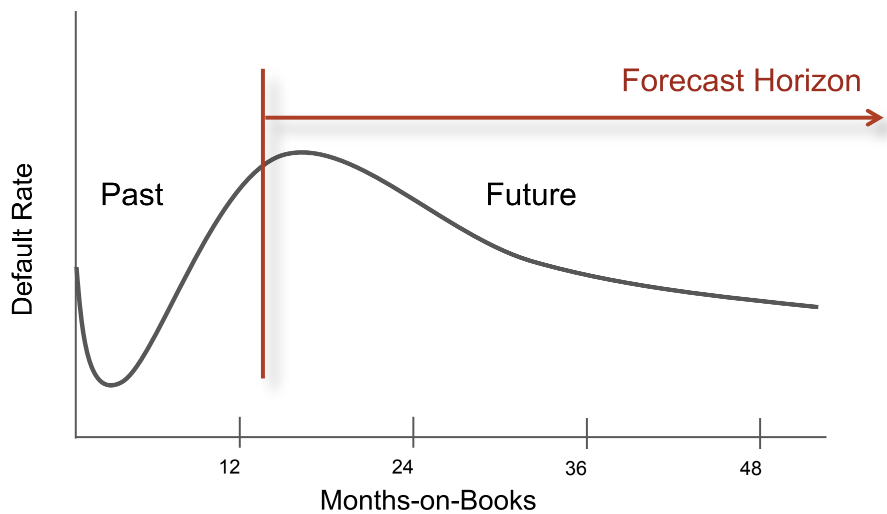


Figure 2: Illustration of reserving for the remaining life of a loan.

Some authors report that a primary motivation for adopting a lifetime loss approach was the difficulty in determining the loss emergence period. The lifetime loss concept is itself useful, providing information to loan pricing and account management. It should also make the loss reserves less volatile through the economic cycle.

Ideally, for a loan of any age, one would like to compute the expected loss for the remaining life until termination. The natural and most well-known method for accomplishing this is vintage analysis. However, vintage analysis is not common at smaller institutions, so the final CECL rules state explicitly that vintage analysis is not required. The models employed in this study all manage to create a remaining life of loan estimation, though only a few explicitly use lifecycles versus age of the loan as is done in vintage analysis.

The examples in the CECL rules also imply that one consequence of not using some form of age-of-loan analysis could be that the same lifetime loss amount is held for the life of the loan regardless of its age. Such an approach is taken in the first example in the final CECL guidelines. However, it assumes that historic data is available that provides an estimate of the full lifetime loss rate for the loan type being considered. For 30-year fixed term mortgages, even with the long history in the current data set, no such direct measure of the 30-year loss rate is available. A model is always required.

Foreseeable Future

CECL distinguishes between the foreseeable and unforeseeable future, Figure 3. For 30-year mortgages being considered in this study, no one can create reliable macroeconomic projections 30 years forward. Instead, most practitioners consider the "foreseeable" future for the economy to be on the order of two years. Beyond this near future, the only plausible economic scenario is to use a through-the-cycle (TTC) average.

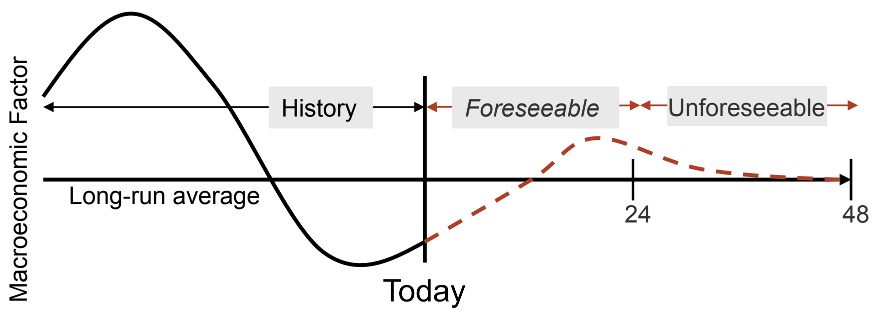


Figure 3: A visualization of a mean-reverting process transitioning between the near-term "foreseeable" future and the long-term "unforeseeable" future.

Some practitioners will create two models, one for the near term with a base macroeconomic scenario and one for the remaining lifetime loss rate, as allowed under CECL. Although this sounds simpler, accurately estimating a TTC loss rate is also non-trivial and the result is two models that must be documented, validated, and audited.

Rather than developing two separate models, the models produced here incorporate mean-reverting macroeconomic scenarios [Breeden & Liang 2015]. This approach provides a smooth transition between the foreseeable and unforeseeable periods. The mean reverting algorithm is referred to as a two-dimensional Ornstein-Uhlenbeck process, which is common in the literature, [Øksendal 2000].

No New Originations or Balance Growth

Under CECL as with FAS 5, the loan loss calculations only apply to loans currently in the portfolio at the time of reporting. Some new loans may appear and charge-off even before the next financial reporting period, but those are outside the ALLL calculation.

In an interesting divergence from IFRS9, CECL considers only the currently outstanding balance. For mortgages this is natural, but for credit cards and other lines of credit, the CECL calculation would not consider any future purchases, only payments toward the currently outstanding balance. Consequently, line of credit calculations will diverge strongly from installment loans under CECL, but those issues will not be considered here.

Aggregate or Loan-level

Comments from FASB about CECL modeling have stated more or less directly that modeling is to be done in aggregate for loan pools that are sufficiently homogenous and loan-level modeling is only to be done for unique loans. These comments often appear to be a mandate for aggregate models, such as vintage models. If true, the author believes that this position will evolve.

Large institutions with CCAR (Comprehensive Capital Analysis and Review) models in most cases will be able to modify them for CECL. The macroeconomic scenarios will need to change, or they may elect to create a second TTC model for the unforeseeable period of the loan. Since the Federal Reserve has been pushing CCAR institutions to develop loan-level models, it seems highly unlikely that FASB would reject loan-level models.

Loan-level methods are invariably more complex, which will weigh in the model selection process for some lenders, but they naturally provide supplementary business value in account pricing and management. Therefore, this study considers aggregate and loan-level methods as equally applicable.

Mortgage Data

A combined, publicly available data from Fannie Mae and Freddie Mac was used for model creation and assessment. This data provides origination and performance information on conforming 30-year fixed rate mortgages.

In addition to monthly loan status, the database contains a number of attributes suitable for loan-level credit risk estimation. The full list of data fields for Fannie Mae and Freddie Mac is given below.

Table 1: Origination and performance data fields available in the Fannie Mae and Freddie Mac datasets.

Origination fields	Performance fields
Loan sequence number	Loan sequence number
Credit score	Monthly reporting period
First payment date	Current actual UPB
First time homebuyer flag	Current loan delinquency status
Maturity date	Loan age
Metropolitan statistical area	Remaining months to legal maturity
Mortgage insurance percentage	Repurchase flag
Number of units	Modification flag
Occupancy status	Zero balance code

CLTV -- cumulative loan to value
DTI -- debt to income
Original UPB (Unpaid balance)
LTV -- loan to value
Original interest rate
Channel
PPM flag
Product type
Property state
Property type
Postal code
Loan purpose
Original loan term
Number of borrowers
Seller name
Servicer name

Zero balance effective date
Current interest rate
Current deferred UPB
DOLPI -- Date of last paid installment
MI recoveries
Net sales proceeds
Non MI recoveries
Expenses

For the models developed in the study, the following definitions were used.

Table 2: The following definitions were available to define key variables for modeling.

Variable	Definition
Default	Current loan delinquency status>=6, i.e. 180+ days past due (DPD)
Active	Non-default and current actual UPB>0
Attrition	Zero balance code=1 (prepaid)
Outstanding balance	Current actual UPB if status = active
Default balance	Current actual UPB if status = default
Origination balance	Current actual UPB if Current Date=Vintage
Loss	Default balance + Accrued Interest + Total Costs – Total Proceeds
Accrued Interest	Default balance*((Current Interest Rate/100-0.0035)/12)*(Months between Last Principal and Interest paid Date and zero balance date)
Total Costs	Foreclosure Costs + Property Preservation and Repair Costs + Asset Recovery Costs + Miscellaneous Holding Expenses and Credits + Associated Taxes for Holding Property
Total Proceeds	Net Sales Proceeds + Credit Enhancement Proceeds + Repurchase Make Whole Proceeds + Other Foreclosure Proceeds

The Fannie Mae and Freddie Mac portfolios represent most but not all of the conforming mortgage industry. The data made available by those institutions is a large share of their portfolios, but not the entirety. The following graphs summarize the volume of loans in the provided data as well as defaults and default rate.

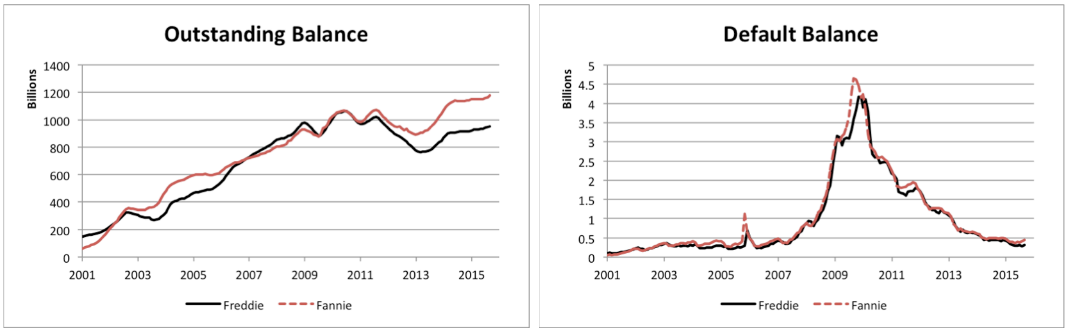


Figure 4: Outstanding balances and monthly default balances for the Fannie Mae and Freddie Mac datasets.

In Figure 5 macroeconomically-driven default peaks are seen in 2002 and 2009. The short spike in 2006 is the bankruptcy law change.

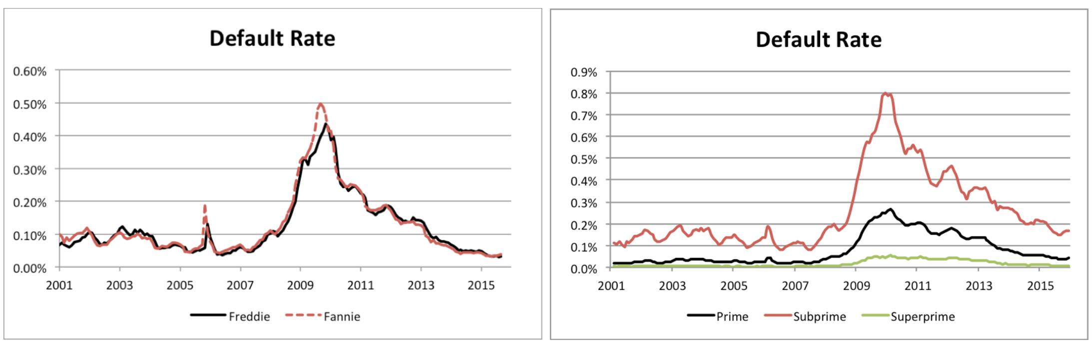


Figure 5: The monthly default balance rate is shown for the two data sources and for the three risk grade segments on the combined data set.

Figure 6 shows the monthly loan origination volume by risk grade segment. The abrupt end of subprime lending is 2008 is clear, as well as the boom period for overall mortgage lending in 2003.

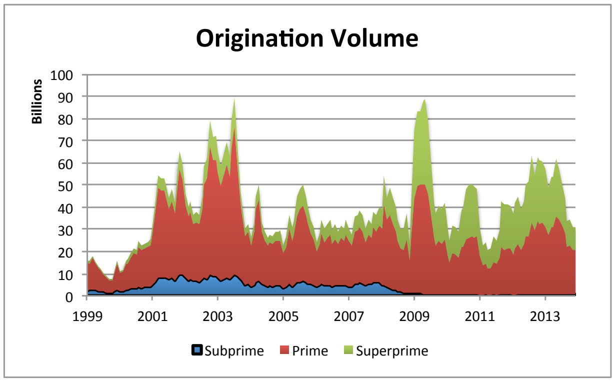


Figure 6: Monthly mortgage origination volume split by FICO segment.

Segmentation

All models were built on national data and segmented by state. Most models also included segmentation by FICO score: Subprime is less than 660, Prime is 660 to 780, and Superprime is 780 and above. Time series of default rates by FICO segment are shown in Figure 5.

Macroeconomic Data

As part of the government's implementation of the Dodd-Frank Stress Test Act (DFAST), the Federal Reserve Board regularly releases Base, Adverse, and Severe scenarios for a set of macroeconomic factors. Our expectation is that many if not most lenders will choose to adopt the DFAST Base scenario for their CECL loss reserve estimation. Since these factors and scenarios have become industry standards, this study has focused on the use of these factors for incorporating macroeconomic sensitivity.

Table 3: Domestic macroeconomic factors available in the DFAST scenarios.

Mortgage-related Factors	Other Factors
Real GDP growth	Nominal GDP growth
Real disposable income growth	Nominal disposable income growth
Unemployment rate	CPI inflation rate
Mortgage rate	3-month Treasury rate
Dow Jones Total Stock Market Index	5-year Treasury yield
House Price Index	10-year Treasury yield
	BBB corporate yield
	Prime rate
	Commercial Real Estate Price Index (Level)
	Market Volatility Index (Level)
	Auto48 rate
	Credit Card rate
	Personal24 rate
	Other Factors
	Nominal GDP growth

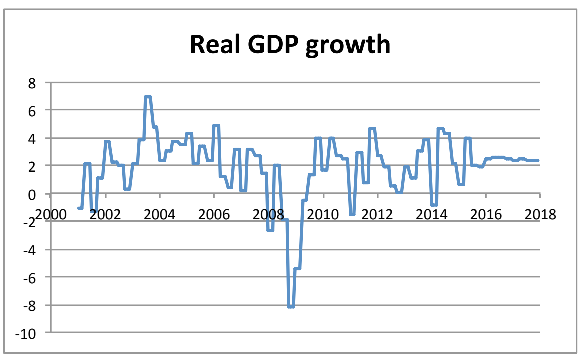
Nominal disposable income growth
CPI inflation rate

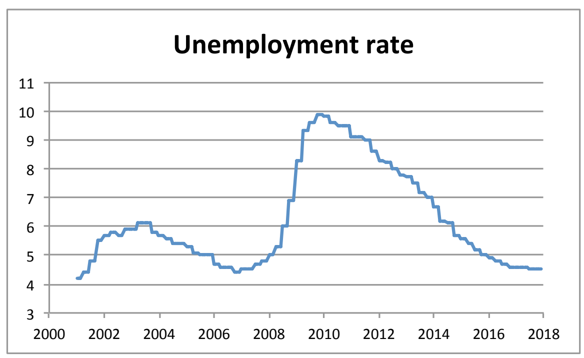
Table 3 lists all of the available domestic macroeconomic factors available in the DFAST scenarios. The left column lists those factors as most naturally related to mortgage performance and therefore considered in this study. The right column lists additional factors that are either redundant with the left column or less directly related to mortgage performance. The list of candidate factors was restricted in this manner to avoid overfitting when estimating the models.

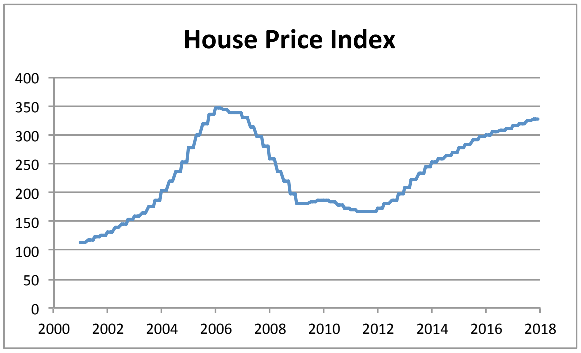
For modeling by geographic segmentation, state-level data for the DFAST factors was obtained from the Federal Reserve Economic Database (FRED). The DFAST national scenarios were apportioned to the states through a set of lead/lag and scaling models. Such models are not advertised as accurate state-level macroeconomic forecasting models, but they are sufficient to provide plausible state-level scenarios that are consistent with the DFAST national scenarios and past macroeconomic sensitivities.

Macroeconomic Scenarios

For forecasts beginning in January 2016, the first two years of the economic scenarios follow the Federal Reserve Board's 2016 DFAST scenarios. Figure 7 shows graphs of the history and scenarios for some of the variables used in creating the macroeconomic models.







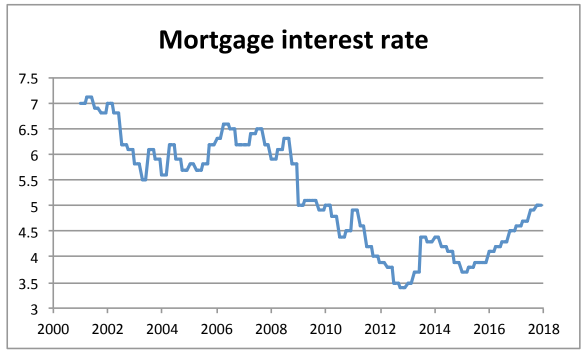


Figure 7: Graphs of historic macroeconomic data through the end of 2015 and the FRB Base scenario thereafter.

Modeling Approaches

The final CECL guidance provides a laundry list of techniques that may be used for estimating loss reserves. Rather than being a list from which practitioners should choose, the list illustrates a philosophy that any approach is allowed as long as it adheres to the guidelines and is appropriate for the size and complexity of the institution (or portfolio).

The US Federal Reserve released a document on best practices in stress testing in 2013 [BGFRS 2013]. That document described the range of practice in creating stress test models along with pros and cons from the Fed's perspective. In many respects, CECL could be viewed as a stress test model run with specific scenarios for macroeconomic drivers, no new originations, and no balance growth, so the Fed's recommendations are also worth consideration.

After reviewing both of these documents, standard industry practice, and the academic literature, we have selected the following set of models as most representative of what lenders are likely to create.

Table 4: Models types tested for lifetime loss estimation under CECL.

Model Type	Level
Moving Average	Risk Segment
Historic Precedent	Risk Segment
Time Series	Risk Segment
Roll Rate	Risk Segment
Vintage	Vintage by Risk Segment
State Transition	Loan-level by Risk Segment
Discrete Time Survival	Loan-level by Risk Segment

One confusing aspect of the CECL guidance is the way discounted cash flows are listed as a model type along side vintage models, roll rates, and others. At first glance, the implication would seem to be that discounted cash flows and other model types are mutually exclusive. However, subsequent remarks appear to clarify that these concepts may be combined. For this study, DCF is viewed as an aggregation method on the outputs of a loss model rather than a modeling technique in itself. DCF will be discussed more in a later section.

All of the models below will predict through default balance. For the final reported totals, an assumption for recovery rate is used, based upon published mortgage industry statistics.

$$Recovery\ Balance(t) = 0.7 * Default\ Balance(t-6)$$

This recovery rate adjustment is done only so that the charge-off balance numbers will be more typical of those in the industry. Obviously a real recovery rate model would be preferable, but that will wait for a later study.

Moving Averages

Averages of past history should not really be considered a modeling technique. Historic average rates only work if everything is steady-state – a constant economy, constant loan growth, and constant origination credit quality. In reality none of these are true, which directly precipitated CECL. Moving averages are included here to provide a comparison of how loan loss reserves have most commonly been computed pre-CECL (not counting qualitative adjustment factors), thus providing a comparison between old and new practices.

Time Series models

The simplest forward-looking model in this study requires creating macroeconomic time series models of the balance default rate and pay-down rate. The default rates are shown in

$$Balance\ Default\ Rate(t) = \frac{Default\ Balance(t)}{Active\ Balance(t-1)}$$

$$Balance\ Payoff\ Rate(t) = \frac{Payoff\ Balance(t)}{Active\ Balance(t-1)}$$

The rates are correlated to macroeconomic factors in order to make scenario-based forecasts. Lifetime losses can then be simulated by projecting these rates under a mean-reverting base macroeconomic scenario until all currently outstanding balances are either paid-off or defaulted.

$$\begin{aligned} Outstanding\ Balance(t) \\ &= Outstanding\ Balance(t-1) - Payoff\ Balance(t) \\ &- Default\ Balance(t) \end{aligned}$$

Transformation of the macroeconomic data and model estimation are primary considerations. When comparing two time series, as in Figure 8, the impacts of X on

Y are generally not instantaneous. For example, six months may be required before an increase in unemployment creates a corresponding rise in default rate. The delay in the impact is called the lag. In addition, it may not be a single value of X from six months ago, but rather an average over a range of values that drives Y. Therefore, an averaging transformation requires a specific window.

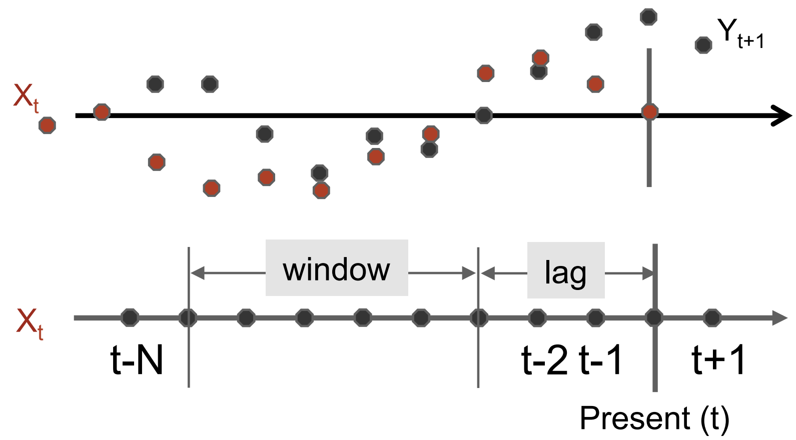


Figure 8: A visual comparison of a performance time series Y and an explanatory variable X.

The notion of lag and window applies to most transformations that can be applied to the economic time series: moving average, difference, and relative change being the most common. The lag and window approach is just a simplification to reduce the number of parameters needed as compared to estimating separate coefficients for each lag as done in Distributed Lag Models [Judge et. al. 1985].

Table 5 lists the variables considered and the transformations tried. Log-ratio is the measure of relative change and refers to the following transformation:

$$\text{log– ratio}(x; w) = log\left(\frac{x(t)}{x(t-w)}\right)$$

For small changes, log-ratio is roughly equal to percentage change, but it has better estimation and extrapolation properties for time series modeling.

Table 5: The list of variables and transformations tested for predictive ability.

Variable	Transformation
Real GDP	Log-ratio
Nominal Disposable Income	Log-ratio
Unemployment rate	Moving average
Unemployment rate	Difference
Unemployment rate	Log-ratio
House Price Index	Log-ratio
Mortgage Interest rate	Difference

DJIA	Log-ratio

To optimize lag and window for each variable, an exhaustive search is performed over a range of parameters. For this study, we allowed lags from 0 to 12 and windows from 1 to 24. The final model selects from among these candidates transformed macroeconomic factors to optimize in-sample fit and satisfy constraints on the allowed sign of the relationship. For example, unemployment rate must be positively correlated to default rate or the result is assumed to be spurious and rejected.

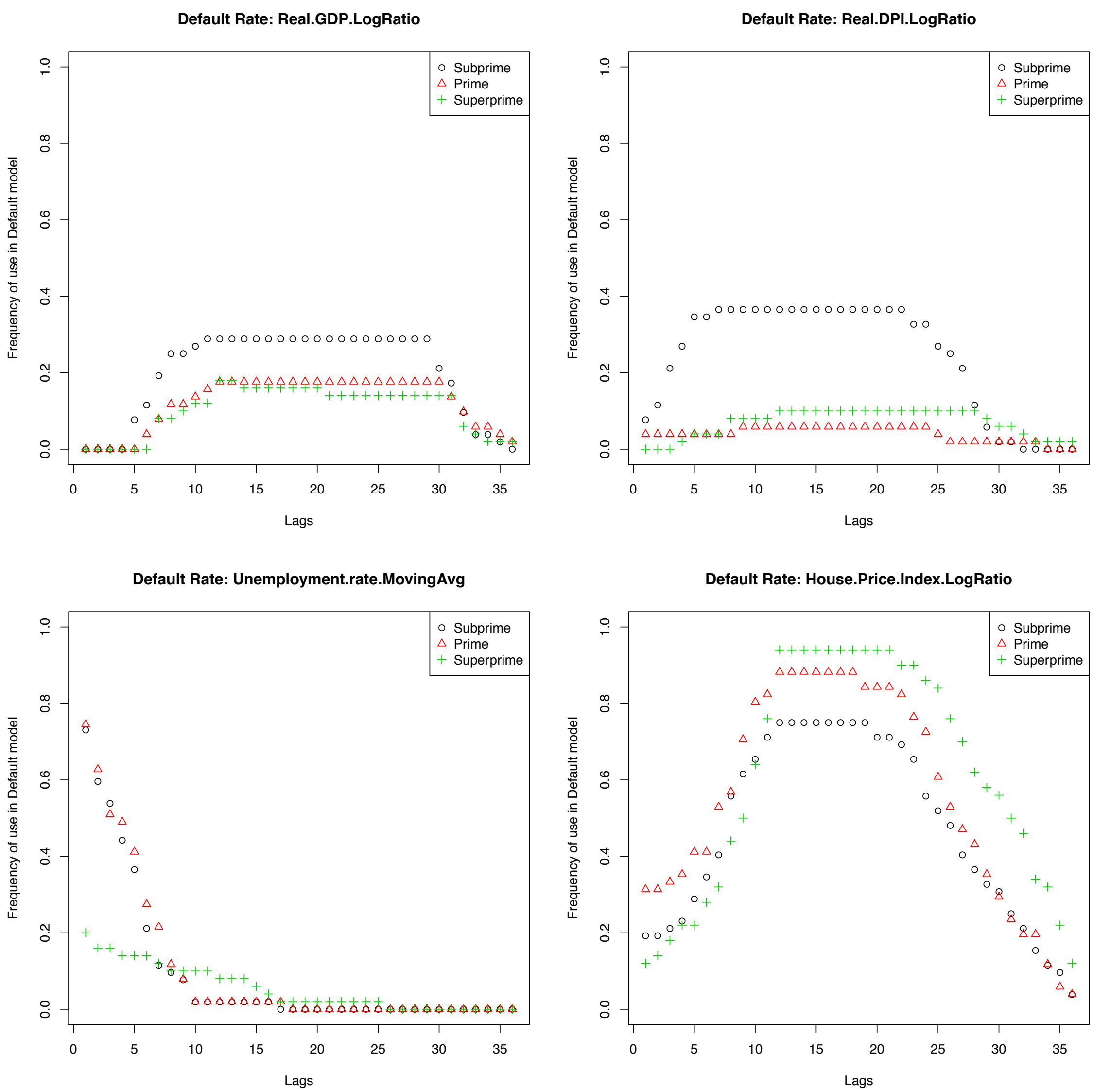
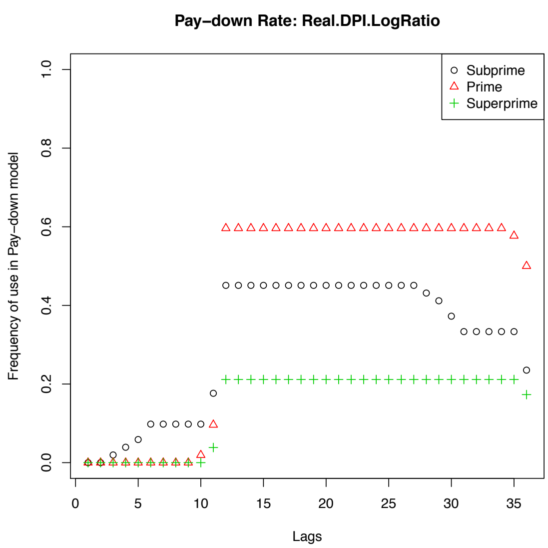
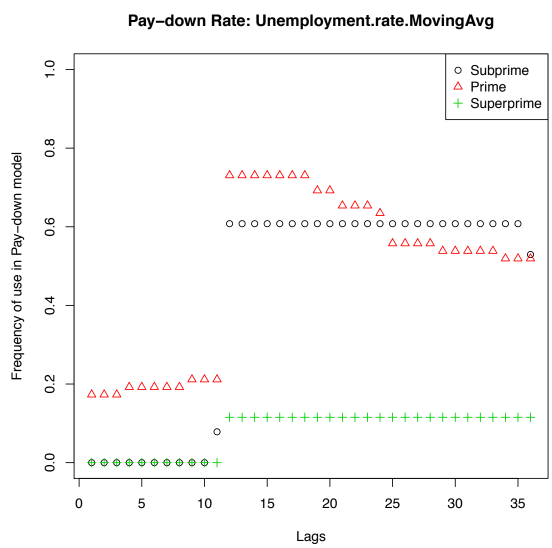


Figure 9: By summarizing across all of the state-level macroeconomic models for balance default rate, the above graphs show the frequency of occurrence for the value at a given lag to appear in the model. The four variables shown are the most prevalent in the models.

Figure 9 and Figure 10 summarize the state-level macroeconomic models. Each model used the window-and-lag structure described above. If a specific lag for a transformed macroeconomic factor was included in the model for a state, it received one count among all states. Therefore, ~0.9 frequency for lags between 12 and 20 for the log-ratio of house price index (Figure 9 bottom right) indicates that >90% of all states included those transformed HPI values in the final model.

The graphs show strong consistency among states, each modeled independently by risk segment. In addition, patterns appear indicating that GDP, DPI, and unemployment are most important to subprime and prime borrowers, whereas HPI is most important to superprime borrowers. This suggests that wages and employment are most important to lower FICO consumers and asset values are most important to higher FICO consumers.





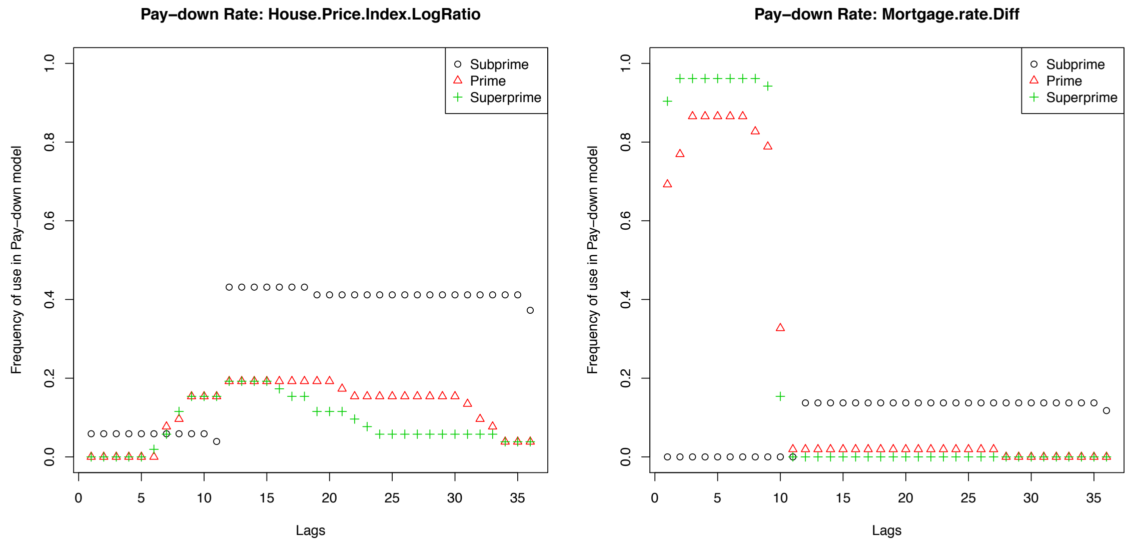


Figure 10: By summarizing across all of the state-level macroeconomic models for balance pay-down rate, the above graphs show the frequency of occurrence for the value at a given lag to appear in the model. The four variables shown are the most prevalent in the models.

The graphs for balance pay-down rate show that income-related variables DPI and unemployment are most important for subprime and prime borrowers, HPI has mid-level importance, but changes in interest rates are most important for superprime and prime borrowers. Subprime borrowers show almost no sensitivity to interest rate changes, presumably because of an inability to refinance.
Pure macroeconomic time series models necessarily assume that all portfolio dynamics are explainable by the economy. Since portfolio management actions are typically correlated to economic conditions, this implies that the lender's future underwriting and account management decisions are predictable from macroeconomic scenarios based upon past actions. For the entire lending industry, previous studies suggest this may be true [Breeden & Canals-Cerdá 2016]. However, this is a strong assumption for an individual lender.

Roll Rate Models

For the last forty years, the two most common kinds of models for retail lending portfolios are credit scores and roll rates. Roll rate models [Coffman et. al. 1983] are similar to a state transition model, but estimated on aggregate net balance flows from one delinquency bucket to the next.

$$R_i = \frac{Balance\ Bucket\ i(t)}{Balance\ Bucket\ i-1(t-1)}$$

The net balance roll rates are visualized in Figure 11.

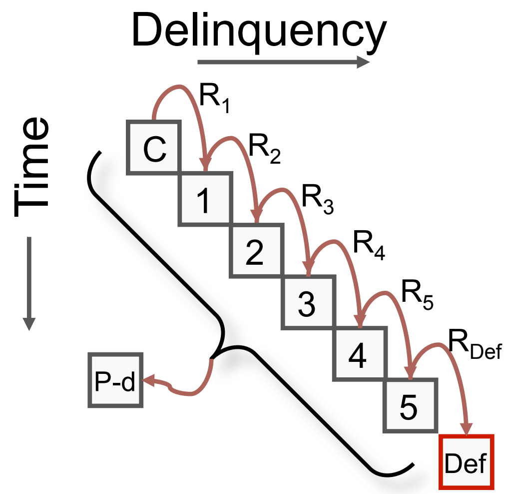


Figure 11: A diagram showing the definitions of the net roll rates through the delinquency buckets and the pay-down rate.

Historically models of roll rates have been simple moving averages of the rates, but for CECL estimation the net balance roll rate is modeled with macroeconomic factors in the same manner described for the time series model. In addition, the balance pay-down rate as defined previously is modeled with macroeconomic data so that both default and pay-off end states are included. Thus, the roll rate model is similar to the time series model, but with intermediate delinquency transitions added.

Figure 12 shows the roll rates time series split by risk grade. Each time series is modeled separately with macroeconomic data. The time series show the expected structure and with strong similarities between risk grades.

The final lifetime loss is calculated by summing the monthly losses until all existing loans reach zero balance, as described for the time series model.

Figure 12: Graphs showing roll rate time series by risk segmentation.

Vintage models

Vintage models [Breeden 2014] naturally capture the timing of losses and attrition versus age of the loan, and therefore are an obvious choice for lifetime loss calculations. Age-Period-Cohort models [Holford 1983, Glenn 2005] are commonly used to estimate such models. The key rates for modeling are probability of default (PD), attrition rate (AR), and exposure at default (EAD). LGD is not being modeled. Instead, recoveries are being held at 70% for all models.

$$PD(t) = \frac{Default\ Accounts(t)}{Active\ Accounts(t-1)}$$

$$AR(t) = \frac{Attrition\ Accounts(t)}{Active\ Accounts(t-1)}$$

$$EAD(t) = \frac{Default\ Balance(t)}{Outstanding\ Balance(t-1)}$$

Each key rate is decomposed into a lifecycle function versus the age of the loan $F(a)$, a credit risk function versus origination date (vintage) $G(v)$, and an environment function versus calendar date $H(t)$. The lifecycle function captures the timing of losses or attrition. The environment function is an index of sensitivity to macroeconomic changes, which is then correlated to macroeconomic factors as was done with the time series and roll rate models.

$$rate(a, v, t) \sim F(a) + G(v) + H(t)$$

The decomposition is performed assuming a binomial distribution for PD and AR and a lognormal distribution for EAD.

Figure 13 shows a visual example of the decomposition process for the subprime, prime, and superprime segments. The figure at left shows default rate time series aggregated by annual vintage. The top right graph is the lifecycle function transformed to the monthly probability of default, $1/1 + \exp\left(-F(a)\right)$. The middle right graph shows the change in log-odds of default by vintage, $G(v)$, where the zero line is the average default rate. The vintage function shows the credit cycle with high risk loans being originated in 2005 through 2008. The bottom right graph measures the change in log-odds of default where the zero line is the average environment, $H(t)$. The environment function is capturing the economic cycle with the onset of the 2009 recession clearly visible.

All three functions were estimated nonparametrically [Schmid & Held 2005], meaning that every monthly point on the functions was an independent parameter. This approach is effective for complex products that cause significant structure in the lifecycles, such as adjustable rate mortgages with teaser rate periods. In the present example, all these lifecycle parameters could be replaced with a few spline coefficients or transformations of account age, as done later in the state transition models.

Macroeconomic scenarios are used to project the future value of the environment function [Breeden et. al. 2008], which is then combined with the vintage and lifecycle functions to produce monthly forecasts for each vintage. Lifetime loss forecast sums across vintages and calendar date to the end of the term or until the outstanding balance reaches zero.

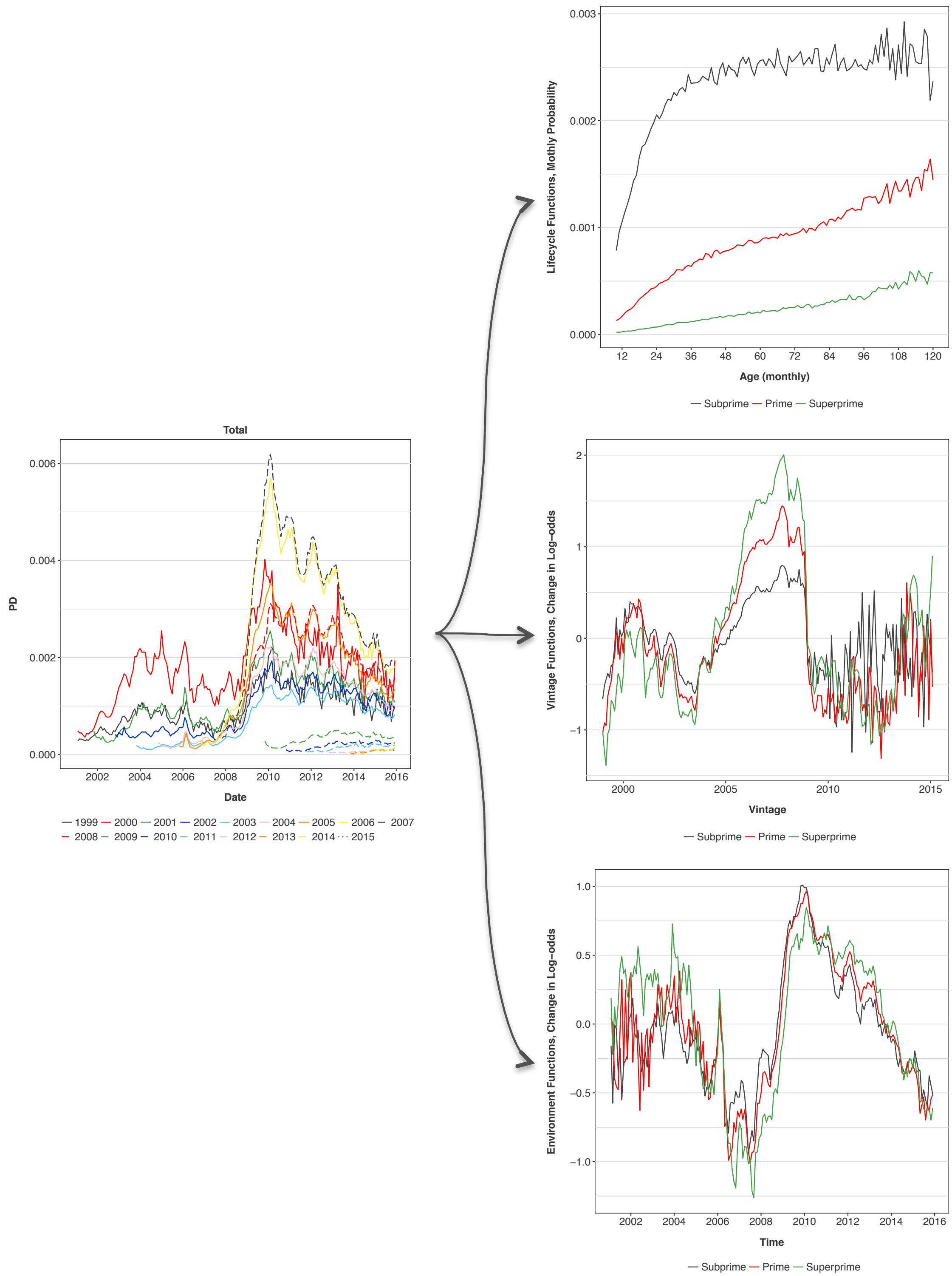


Figure 13: A visualization of the decomposition process. Vintage time series (left graph) is decomposed into three functions. Upper right shows the monthly probability of default by age of the loan. The middle right graph measures the change in log-odds of default versus vintage, capturing the credit cycle. The lower right graph measures the change in log-odds of default versus calendar date, capturing the economic cycle. The decomposition results are measured by FICO segment.

State Transition models

State transition models [Thomas et. al. 2001, Bangia et. al. 2002] are the loan-level equivalent of roll rate models. They derive from Markov models, though in practice they may not satisfy the Markov criteria.

Rather than modeling aggregate movements between delinquency states, the probability of transition is computed for each account. The states considered are current, delinquent up to a maximum of six months delinquent, default, and pay-off. Account transition probabilities are modeled rather than the dollar transitions in the roll rate model.

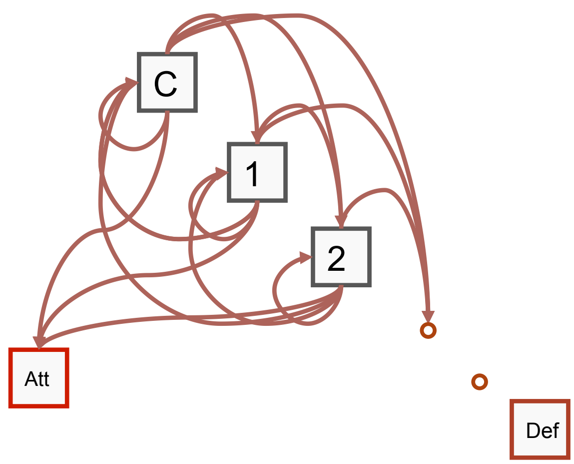


Figure 14: A visualization of the possible transitions between the various delinquency states. Only the transitions for the first few states are shown.

Not all transitions are populated well enough to be modeled, even with the large dataset available. Table 6 shows the transitions modeled with regression models, the ones set to constant rates, and the ones set to zero.

Table 6: Average transition probabilities between all possible states. The green transitions are fully modeled. The yellow transitions are held as constants. The white transitions were too thin to model effectively.

Pr(i->j) %	Current	1m Delq	2m Delq	3m Delq	4m Delq	5m Delq	Default	Attrite
Current	97.38%	0.88%	0.01%	0.00%	0.00%	0.00%	0.00%	1.73%
1m Delq	36.97%	44.88%	16.67%	0.09%	0.02%	0.00%	0.01%	1.35%
2m Delq	12.57%	16.48%	34.31%	35.44%	0.20%	0.03%	0.02%	0.90%
3m Delq	7.08%	3.64%	8.01%	20.45%	59.53%	0.21%	0.08%	0.83%
4m Delq	7.13%	1.14%	1.38%	3.49%	15.32%	69.88%	0.28%	0.80%
5m Delq	6.71%	0.78%	0.51%	0.78%	2.35%	12.46%	74.26%	0.77%

The transitions with enough data are modeled via logistic regression as monthly probabilities dependent upon macroeconomic factors, loan level factors, and transformations of the age of the loan, namely $age, age^2, \log(age), log^2(age)$. When the model for a transition is created, if some factors have insignificant coefficients, the least significant is removed and the estimation is repeated until all factors are significant. A binned factor is considered significant so long as it has a significant bin, because the zero level for the bins, and thus the reference point for the p-value estimation, is arbitrary.

Note that multinomial logistic regression is a preferable approach to estimating a set of independent logistic regression models for the transition. However, multinomial logistic regression required so much memory (estimating many transitions simultaneously) that the data set had to be reduced to the point where some of the transitions became unstable for modeling. The above approach is therefore a compromise.

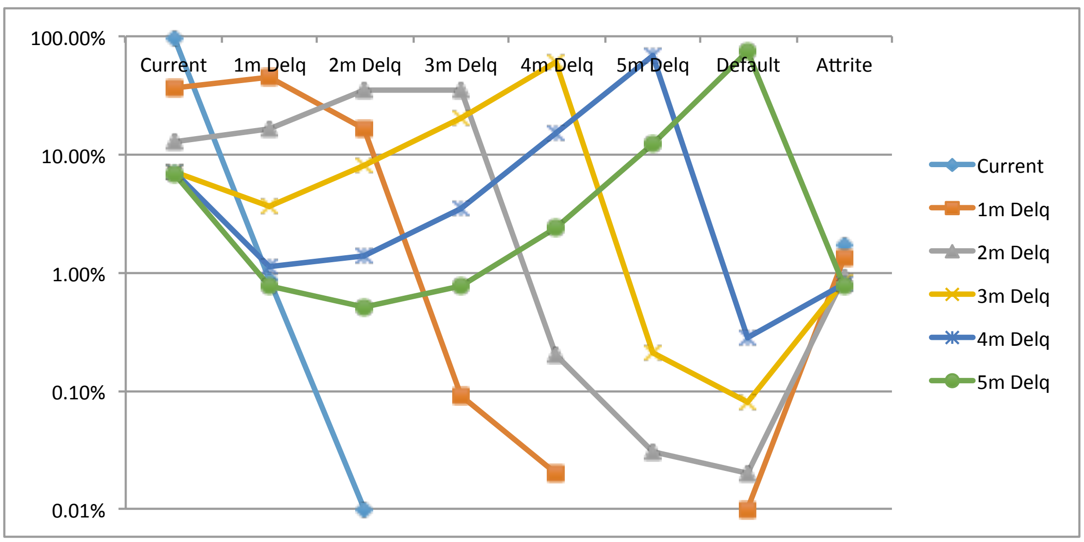


Figure 15: A graph of the transitions in Table 6 from each active state to all possible future states. The lines are the active states and the points are transitions to future states.

The complete transition modeling approach used here is possible because of the large data set, but also results in a large number of parameters. Practitioners often simplify the process by making assumptions about relationships between the transitions. As seen in Figure 15, the transitions do show a pattern, but not that of a simple distribution. For the current study, no further simplifications were explored, but the structure seen would justify such efforts.

The balances are predicted as

$$
\begin{aligned}
Oustanding\ Balance(t) \\
&= Outstanding\ Balance(t-1) - Scheduled\ Principal\ Payment(t) \\
&\quad - \ Surplus\ Principal\ Payment(t)
\end{aligned}
$$

The surplus principal payment is obtained by modeling

$$Surplus\ Principal\ Payment(t) = \frac{Surplus\ Principal\ Payment(t)}{Scheduled\ Principal\ Payment(t)}$$

This variable is modeled the same as the state transitions, but with the assumption of a lognormal distribution.

Discrete Time Survival models

Survival models [Cox & Oakes 1984] are related to vintage models, but usually with the implication of creating loan-level models with scoring attributes. Cox proportional hazards [Cox 1972] models are the original and classic approach to creating such models, but they were developed with continuous time in mind. For monthly sampled data such as available here and in lending in general, discrete time survival models are employed.

Once the change to discrete time is made, the result is just a logistic regression panel model of PD or AR. Practitioners commonly just include age as a factor in the regression, either nonparametrically or as a set of transformations as shown in the state transition model. Transformations of macroeconomic factors and scoring factors are included as well. Then a regression solve is performed.

Because of problems with multicolinearity [Breeden & Thomas 2016] between age of the account, macroeconomic factors, and behavioral scoring factors, this study follows a modified approach developed by Breeden [2016] sometimes referred to as APC Scoring. Namely, the lifecycle and environment functions from the vintage model (APC) estimation are used as fixed inputs to a panel data model with scoring attributes. This concentrates the population dynamics into the lifecycle and environment functions and leaves the scoring factors as loan-level idiosyncratic effects, thereby solving the multicolinearity problem and eliminating any need to forecast the behavioral factors.

$$log\left(\frac{p_i(a,v,t)}{1-p_i(a,v,t)}\right) = F(a) + H(t) + c_0 + \sum_{j=1}^{n_s} c_j s_{ij} + \sum_{v=1}^{n_v} g_v$$

Separate origination and behavioral models are built, the former using only factors available at origination and the latter using both origination factors and behavioral factors such as recent delinquency. For the behavioral models, the coefficients are a function of forecast horizon, because any delinquent account will have either cured or defaulted within six to twelve months. The remainder of the forecast is thus dominated by persistent factors like FICO score and LTV.

Loan-level estimates of PD and AR were created by the process above. EAD followed the same process, but with a lognormal distribution.

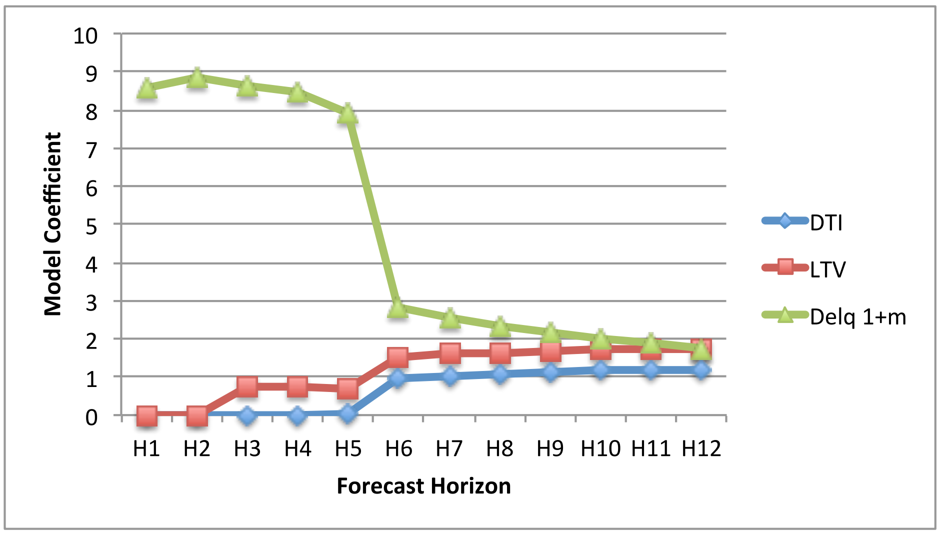


Figure 16: Coefficients are shown for three of the factors in the Prime segment PD behavior model. The coefficients are a function of forecast horizon. Forecasts beyond 12 months use the 12-month coefficients.

The final lifetime loss forecasts are created by aggregating the loan-level monthly loss estimates.

Discounted cash flows

Computing discounted cash flows (DCF) is a standard technique in finance for considering the time value of money. A payment (or loss) in five years does not have the same importance to the lender as a loss in five months. Discounting the monthly values by the interest rate makes this adjustment. For CECL, the effective interest rate on the loan is to be used for discounting. With fixed term loans, the effective interest rate is just the rate on the loan. For line-of-credit products it may be a more complex question.

If a model is used to compute the expected loss amount each month, then the formula for computing the discounted lifetime loss is given below

$$\text{Discounted Lifetime Loss} = \sum_{i=0}^{N} \frac{\text{Loss}(i)}{(1 + r_{\text{eff}}/12)^{i}}$$

where r_{eff} is the effective annual interest rate of the loan, N is the number of months in the forecast horizon, and $\text{Loss}(i)$ is the expected loss amount predicted for a specific month. Any model that produces estimates at intervals may be discounted.

The effective interest rate is computed as

$$r_{\text{eff}} = \left(1 + \frac{r}{n}\right)^n - 1$$

where r is the nominal annual interest rate and n is the number of compounding periods per year. For mortgages this is typically monthly, $n = 12$.

Although DCF is common practice, not all lifetime loss estimates may produce monthly loss values. To accommodate such situations, the CECL guidelines again state that discounted cash flows are not required. However, without DCF the lifetime loss reserves may be substantially higher than with DCF.

Model Comparisons

After the models were estimated, they were run through a variety of tests to assess their properties and impacts for CECL.

In-sample Accuracy

The first question for any model is how accurately it predicts.

Model	Jan 2007 – Dec 2009	Jul 2010 – Jun 2013	Jan 2012 – Dec 2014	Avg Absolute Error
Historic Average	-69.1%	54.1%	65.7%	63.0%
Historic Average by geography	-70.4%	54.3%	62.7%	62.4%
Time Series	11.2%	-28.7%	-12.5%	17.4%
Time Series by geography	19.4%	-26.1%	-12.9%	19.4%
Roll Rate	27.0%	-25.0%	-11.7%	21.2%
Roll Rate by geography	25.8%	-16.7%	-4.5%	15.7%
Vintage	3.6%	3.3%	1.9%	2.9%
Vintage by geography	-2.4%	1.2%	1.5%	1.3%
State Transition	7.8%	11.1%	-1.3%	6.7%

State Transition by geography	-6.2%	12.5%	0.0%	6.3%
Discrete Time Survival	-0.5%	4.5%	3.5%	2.8%
Discrete Time Survival by geography	-3.8%	4.2%	2.9%	3.6%

Table 7 shows the cumulative error over a three year, in-sample forecast starting at three different points in the history. The starting points were selected to be just before the onset of the recession, at the peak of the recession, and during the recovery.

Model	**Jan 2007 – Dec 2009**	**Jul 2010 – Jun 2013**	**Jan 2012 – Dec 2014**	**Avg Absolute Error**
Historic Average	-69.1%	54.1%	65.7%	63.0%
Historic Average by geography	-70.4%	54.3%	62.7%	62.4%
Time Series	11.2%	-28.7%	-12.5%	17.4%
Time Series by geography	19.4%	-26.1%	-12.9%	19.4%
Roll Rate	27.0%	-25.0%	-11.7%	21.2%
Roll Rate by geography	25.8%	-16.7%	-4.5%	15.7%
Vintage	3.6%	3.3%	1.9%	2.9%
Vintage by geography	-2.4%	1.2%	1.5%	1.3%
State Transition	7.8%	11.1%	-1.3%	6.7%
State Transition by geography	-6.2%	12.5%	0.0%	6.3%
Discrete Time Survival	-0.5%	4.5%	3.5%	2.8%
Discrete Time Survival by geography	-3.8%	4.2%	2.9%	3.6%

Table 7: The cumulative percentage error is shown over a three-year forecast period using actual economic history.

The moving average model, which computes the average balance default and pay-down rates of the previous twelve months, is shown only for comparison to common practice before CECL. As a model, moving averages are always out-of-phase with the economic cycle – the reason CECL was created.

Time series and roll rate models perform reasonably well, because they can capture the economic cycle, but not the credit cycle. Roll rate models can be expected to perform reasonably well for the first six to twelve months (better than a time series model) but that advantage disappears for long-lived forecasts such as are needed for 30-year mortgage. Such models may be usable depending upon the model acceptance criteria and the available data.

Vintage models incorporate variation in credit cycles and economic cycles, so their improved accuracy over roll rate and time series models for long-term forecasting is not a surprise. Loan-level models (State Transition and Discrete Time Survival) are more actionable than the preceding aggregate models, but not more accurate than vintage models when viewed in aggregate for the portfolio.

Segmenting by geography (US states) significantly increases the effort, but not the accuracy of the models. Of course, a state-level model will be more accurate than a national model when viewing only one state, but in aggregate state-level analysis does not improve the result. Again, state-level analysis will be more actionable than a single national model, especially for lenders with a limited geographic footprint, but such benefits will not be seen in aggregate for a geographically diversified lender.

Few standards exist in model validation. No fixed threshold exists for how accurate is accurate enough. We can assume that the moving average model is not sufficient, but the other models require further review.

At the other extreme, the author has seen validation teams that require any accepted model to be within 2% error for a test such as this, or some similar criterion. From experience, we know such thresholds are rarely obtainable. Most often, modeling teams reach these objectives only through extensive over-fitting, inclusion of error correcting terms, or other tricks that do not capture the true out-of-sample performance of the model. With the models built here, no such tricks were incorporated. Each model was refined according to the usual practices of passing p-value thresholds on factors, multicolinearity tests, etc. The model structures were refined to provide a fair representation of the technique, but no hand-tuning was performed or other attempts purely to enhance in-sample accuracy. Given this process, we consider the accuracies given here to be representative of what should be expected for products and datasets such as this.

Error Scaling by Volume

The state-level models provide an easy opportunity to test the scaling properties of the models to smaller data sets. Each point along the lines in Figure 17 represents the accuracy of the model for one state. The x-axis is the dollar volume of defaults in the training data. The y-axis is the error on a log scale, so that a random model would be at 0.0 and a model with 10% cumulative error over the three year forecast would be at -1. Three lines are shown for each model corresponding to the three risk grade segments.

Several patterns are apparent in the error scaling with volume, Figure 17. The time series model is relatively insensitive to the volume of training data, appearing as flat lines for the subprime, prime, and superprime segments. However, even adjusting for the volume of the training data, the subprime segment is easier to model with macroeconomic data than the prime segment, which in turn models better than the superprime segment.

The other models are more dynamic with the amount of training data, trending to higher error for small data sets but hitting an accuracy floor for larger data sets. Although all models show a similar pattern, the roll rate model consistently has the highest error. Although roll rates are assumed to be accurate over the first six months, those benefits do not persist through a three-year test such as this.
For the subprime and prime segments, the vintage model is the consistent winner. This comes from the ability to quantify the net credit risk by vintage, regardless of the source of that risk. Only for the superprime segment does the loan-level survival model surpass the vintage model, presumably because the estimation of the vintage credit risk becomes noisier on fewer events. The state-transition model is consistently better than the roll rate model but weaker than the vintage and survival models.

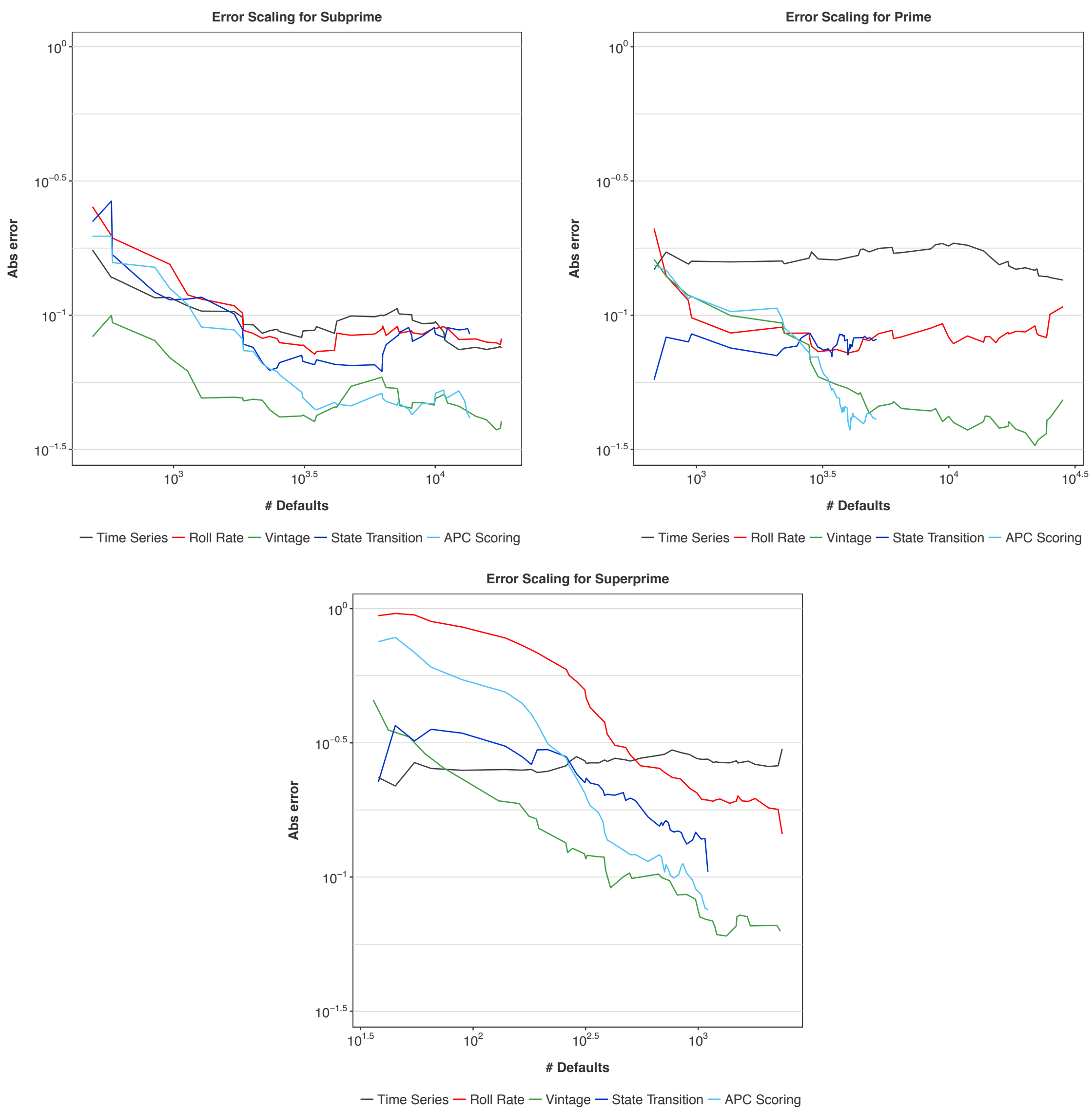


Figure 17: A comparison of model in-sample accuracy versus the volume of charge-off data in the training set. For each model, three lines are obtained corresponding to the subprime, prime, and superprime segments.

Note again that all of these results are in-sample. Out-of-sample tests would have to consider what could actually be known about future economic conditions, whether the coefficient estimates were robust, etc.

Complexity

Accuracy often comes at a price in complexity. To measure model complexity, the following tables simply sum the total number of coefficients required for each model. This does not capture the complexity in software required, but we have few meaningful metrics for that.

Table 8: The total number of estimated coefficients is shown for each national model.

Models, national	Total # of coefficients
Time Series	33
Roll Rate	129
Vintage	69
State Transition	2,097
Discrete Time Survival	240

The vintage model has fewer coefficients than might be expected, because the nonparametric lifecycles are simple enough to replace with a small number of spline coefficients and the environment function was replaced with a simple macroeconomic model.

The state transition model has the largest number of coefficients, because of all the separate transitions being modeled. Various simplifications on state transition modeling exist, all designed to lower the number of estimated coefficients required, but generally at a cost of some in-sample accuracy.

Table 9: The total number of estimated coefficients is shown for each state-level model.

Models, by geography	Total # of coefficients
Time Series by geography	2,159
Roll Rate by geography	6,265
Vintage by geography	2,628
State Transition by geography	75,455
Discrete Time Survival by geography	11,551

The state-level models just scale up the number of coefficients for the 50 independent states being modeled. This could be simplified greatly by including panel model aspects in the design. Having fully independent models is used as an upper bound on the complexity and accuracy for all these approaches.

Computation Times

The models require different amounts of computation time. Times given below are to run one forecast on one 2.5 GHz processor of an AWS server with 240 GB of RAM.

Significant effort was made to optimize the code for the State Transition and Discrete Time Survival models.

Table 10: The time needed to run a single forecast for each model is shown below.

Model	Computation time (min)
Time Series	3
Time Series by geography	9
Roll Rate	3
Roll Rate by geography	10
Vintage	10
Vintage by geography	35
State Transition	1110
State Transition by geography	1680
Discrete Time Survival	270
Discrete Time Survival by geography	390

Overall, the computation time scales with the level of aggregation. Time series to roll rate to vintage is a smooth increase in computation time. Discrete time survival and state transition models are applied at the loan-level, with many more variables being estimated for state transition, so the table affords no real surprises.

Lifetime Forecast Values

The preceding models were used with two-year actual economic values, reverting onto long-run average levels to predict the pool at the forecast start date until then end of term or zero balance. All estimates assume 70% recovery rates. No discounting has been applied. Table 11 shows these lifetime loss forecasts across the models being tested.

Table 11: The CECL lifetime loss estimates are shown starting from different historic points. For historic estimates, real economic data was used for the first two years with mean reversion thereafter. For the January 2016 estimate, the FRB baseline scenario was used for the first two years.

Model	Jan 2007	Jul 2010	Jan 2012	Jan 2016
Historic Average	0.82%	7.14%	2.77%	1.03%
Historic Average by geography	0.73%	6.66%	2.81%	1.09%

Time Series	1.14%	1.71%	1.02%	0.78%
Time Series by geography	1.75%	1.90%	1.23%	1.03%
Roll Rate	1.84%	2.00%	1.28%	1.19%
Roll Rate by geography	2.06%	2.15%	1.39%	1.10%
Vintage	1.52%	2.38%	1.62%	0.90%
Vintage by geography	1.50%	2.38%	1.64%	0.91%
State Transition	1.91%	2.63%	2.18%	1.64%
State Transition by geography	1.51%	2.26%	1.77%	1.08%
Discrete Time Survival	1.47%	2.34%	1.59%	0.63%
Discrete Time Survival by geography	1.38%	2.31%	1.59%	0.68%
Actual*	~2.02%	> 1.8%		

Discounted Cash Flows

The options given for computing loss reserves under CECL include discounted cash flows. Discounting as an economics concept seems used when considering a lifetime loss calculation. Distant future losses or payments should be less important than near-term losses or payments when considering the time value of money.

This concept could be applied to the current context in two ways:

1. Directly discounting the loss estimates. The monthly expected loss forecasts could be discounted and then summed to a total discounted loss.
2. Discounting the payment stream. The existing forecasts of attrition and charge-off are used to adjust the likelihood of receiving scheduled payments. These adjusted payment forecasts are then discounted, summed, and compared to the outstanding balance. This is the approach accountants think of when discussing Discounted Cash Flows (DCF).

For computing discounted losses, the standard discounting formula is applied

$$Discounted\ Loss = \sum_{i=1}^{n} \frac{Loss(i)}{(1+r)^{i-1}}$$

where r is the interest rate of the mortgage and n is the total life of the loan, equal to 360 months in this study. $Loss(i)$ is computed from the forecasting model. Any purchase adjustments to the rate are ignored in this study, since no such information is available.

For computing discounted cash flows, the final loss is calculated as

$$DCF\ Loss = Oustanding\ Balance(k) - PV$$
$$PV = \sum_{i=k+1}^{n} \frac{Total\ Payment(i)}{(1+r)^{i-1}}$$

In this formula, k is the current age of the loan prior to forecasting and $Total\ Payment(i)$ is defined as

$$\begin{aligned} Total\ Payment(i) \\ &= Scheduled\ Payment(i) * \big(1 - adj(i)\big) + Payoff(i) \\ &+ Surplus\ Prepay(i) + Recoveries(i) \end{aligned}$$

where

$$adj(i) = \sum_{j=1}^{i-1} Attrition\ prob(j) + \sum_{j=1}^{i-1} Charge\text{–}off\ prob(j)$$
$$Payoff(i) = Outstanding\ Balance(i-1) * Attrition\ prob(i)$$

$Attrition\ prob$ and $Charge\text{–}off\ prob$ are the probabilities of attrition and charge-off in a given month, conditioned on surviving to that point. These values come from the forecasting model.

Surplus prepayment is a payment beyond the scheduled principal payment, but is insufficient to payoff the loan. This distinction is made because many forecasting models consider loan payoff (attrition) separately from partial balance prepayment. Surplus prepayment also comes from the forecasting model.

This information can then be used to update the outstanding balance.

$$\begin{aligned} Outstanding\ Balance(i) \\ &= Outstanding\ Balance(i-1) - Payoff(i) - Loss(i) \\ &- Principal\ Payment(i) * (1 - adj(i)) \end{aligned}$$

Recoveries in this study are assumed to be 70% of the balance charged-off six months prior. This is a rough approximation, but recovery modeling is not a target of the current study.

Table 12: For the vintage model, a comparison is shown for the simple cumulative loss forecast, discounted losses, and discounted cash flow. The columns indicate different starting points for the lifetime loss forecast.

Calculation	Jan-07	Jul-10	Jan-12	Jan-16
Simple Cumulative Loss	1.52%	2.38%	1.62%	
Discounted Losses	1.08%	1.91%	1.28%	
Ratio to Simple	0.71	0.8	0.79	0.77
Discounted Cash Flow	1.28%	1.96%	1.40%	
Ratio to Simple	0.84	0.82	0.87	0.84

The results in Table 12 show that directly discounting the loss stream would lower the lifetime loss reserve by 20 to 30% depending upon the starting point for the calculation. The discounted cash flow approach results in equivalent reductions by estimating payments rather than losses.

From a theoretical perspective, we cannot say which approach is more appropriate. Discounting cash flows is more common among accountants, but seemingly because they need a net present value calculation for loan pricing. In the present context that starting point may not be necessary. Nevertheless, lenders may prefer the lower numbers and CECL provides justification.

The DCF calculations here do not include other income sources such as late fees. For mortgage lending, those are generally viewed as minor in comparison to interest income. However for line of credit products such as credit cards, penalty fees are a substantial portion of the cash flows. A DCF approach to line of credit products would probably need to include modeling of other fees.

FAS 5 vs. CECL

One of the primary considerations for CECL adoption is the magnitude of the change from the old ALLL calculations under FAS5 to the new rules under CECL. Historically most lenders below roughly $10 billion in assets have used simple moving averages of their loss rates to set their baseline ALLL numbers before adding manual adjustments. Therefore, the following comparison assumes that an average loss rate for the previous 12 months was used with a 24 month loss emergence period to set loss reserves under FAS 5. The lifetime loss numbers from the vintage model are used as the CECL benchmark.

Table 13: The comparison is shown for a simple moving average approach under FAS 5 with a 24 month loss emergence period and a lifetime loss calculation under CECL using the vintage model with direct loss forecasts and discounted cash flows.

Increase in loss reserves under CECL's lifetime loss calculation					
	Segment	Jan-07	Jul-10	Jan-12	Through-the-Cycle Avg
Lifetime Loss	Subprime	512%	73%	49%	211%
	Prime	1219%	95%	39%	451%
	Superprime	1933%	160%	27%	707%
	Total	896%	91%	41%	343%
DCF	Subprime	423%	42%	32%	166%
	Prime	1032%	61%	21%	372%
	Superprime	1019%	92%	-14%	365%
	Total	742%	57%	22%	274%

With direct loss forecasting, the increases are dramatic, because lifetime losses for 30-year mortgage are much higher than 24-month losses. The increases are highest going into a recession, because the old models reserve at pre-recession levels. Levels are high in better-than-average economic environments because of reverting to through-the-cycle levels after two years. Loss reserves are roughly equal during improving economic conditions, because the old models over-reserve during this time.

With the discounted cash flow approach, the same patterns exist through the economic cycle, but with somewhat lower estimates. Therefore, the DCF approach has an overall average increase of 274% whereas the average increase with direct aggregation of lifetime loss forecasts is 343%.

Conclusions

The DFA CECL Study provides a number of interesting results one might use when implementing models for CECL. However, we cannot recommend a one-size-fits-all answer. Rather, the following results are intended to be used to assess trade-offs in CECL implementation details.

Foreseeable Future

Using mean-reverting scenarios here allowed the model to adapt to the current portfolio for the lifetime estimation rather than use an average over past portfolios, but at greater complexity. Conversely, it requires only one model rather than two. Even though most practitioners will use a through-the-cycle average default rate as the long-run model, we know from Basel II that these are actually models with their own complexities in estimation.

Accuracy

Projecting losses via time series models of default and pay-down rates produced an average 3-year cumulative error rate of 17-19%. In itself, that will raise concerns with validators, but the accuracy is unchanging relative to the amount of training data, which can be useful for very small or noisy data sets. Vintage models were consistently high performers in terms of accuracy with 1% to 3% error rates. Discrete time survival models and state transition models both perform well (6.5% to 7.5%), but not better than vintage models, showing that loan-level modeling does not guarantee more accuracy. Vintage, state transition, and survival models all had similar scaling properties versus size of training data. Roll rate models were consistently the worst performers at 15% to 20% error rates. Moving averages of historic loss rates are unsuited to lifetime loss forecasting at 60+% error rates. Overall, roll rate and historic average models should not be used for long-lived products.

Creating separate models by US state did not provide greater accuracy when compared to a single national model of the same portfolio. Geographic segmentation provides advantages in business application but not model accuracy.

The guidelines state that vintage modeling is not a requirement. If we assume that "vintage model" refers to any approach that adjusts credit risk and prepayment risk based upon the age of the loan, then the results show significant increases in accuracy for techniques incorporating this (vintage models, state-transition models, and discrete time survival models) as compared to those that do not include it (time series and roll rates).

Accuracy vs. Complexity

The loan-level models (state transition and survival) were by far the most complex in terms of numbers of coefficients and computational time. This complexity did not provide any increased accuracy relative to vintage models, but it does provide business value in account management, collections, pricing, and strategic planning.

The added complexity of roll rate models when compared to time series models provided little benefit other than the change to be more accurate for the first six months of the forecast. Vintage models were the overall winners in the accuracy versus complexity trade-off, so long as sufficient data exists for robust estimation.

Optional DCF

Starting from a lifetime loss forecast, using a time-value of money discounting of the projected monthly losses using the par rate on the mortgage results in a 20% to 30% decrease in the reserve amount. Estimating the principle and interest payments adjusted for the risk of default or prepayment from the loss model and then discounting with the par rate on the mortgage results in an equivalent reduction in the loss reserve as compared to the original lifetime loss forecast.

Old vs. New Rules

The magnitude of the change from the old loan loss rules to CECL will depend strongly on the lifetime of the asset and the point in the economic cycle when the adoption occurs. For 30-year fixed mortgage, the average life of loan is about 5.5 years and the lifetime loss reserve will be 4 times a historic average approach with 24 month loss emergence period. If adoption had occurred just before the onset of the last recession, the adjustment would have been 10x. At the peak of the recession the change would have been 2x. Well into recovery they would have been at parity.

Variability

By design, the new CECL rules provide a significant amount of flexibility in implementation. As seen from this study, even with a straightforward product like 30-year fixed rate conforming mortgages, the range of models listed in the CECL guidelines can produce a range of lifetime loss numbers that vary by a factor of 2. With the option of discounted cash flows, then the range of final answers would vary by more than a factor of 2.

Table 14: A summary of model performance where green is good, red is poor, yellow is in between. Arrows indicate high and low values for the CECL estimates or volatility through the economic cycle.

	Historic Average	Time Series	Roll Rate	Vintage	State Transition	APC Score
Accuracy, 3yr	🔴	🔴	🔴	🟢	🟢	🟢
Accuracy, 6m	🔴	🔴	🟢	🔴	🟢	🟢
Robustness	🟢	🟢	🔴	🟢	🟢	🟢
Complexity	🟢	🟢	🟢	🟡	🔴	🔴
Computation	🟢	🟢	🟢	🟢	🔴	🟡
CECL Estimate	↑	↓	–	–	↑	–
CECL Volatility	↑	↓	↓	↑	–	↑

Being able to choose options that will create such different answers will put the burden on lenders not only to choose the most appropriate models for their portfolios, but in doing so to also choose the level of loss via the models chosen, and to defend that choice to validators, auditors, and examiners.

References

1. Bangia, A., F.X. Diebold, A. Kronimus, C. Schagen, and T. Schuermann (2002), “Ratings migration and the business cycle, with application to credit portfolio stress testing”, *Journal of Banking & Finance*, **26**:2–3, March, pp. 445-474.
2. Board of Governors of the Federal Reserve System (2013) “Capital Planning at Large Bank Holding Companies: Supervisory Expectations and Range of Current Practice”.
3. Breeden, J.L. (2014), "Reinventing Retail Lending Analytics: Forecasting, Stress Testing, Capital, and Scoring for a World of Crises – Second Impression", Riskbooks.
4. Breeden, J.L. and Sisi Liang (2015), “A mean-reverting model to create macroeconomic scenarios for credit risk models”, *Journal of Risk Model Validation* ***9****(4), pp. 1–12*.
5. Breeden, J.L. (2016), “Incorporating lifecycle and environment in loan-level forecasts and stress tests”, *European Journal of Operational Research*, **255**(2), pp. 649-658.
6. Breeden, J.L. and J.J. Canals-Cerdá (2016), “Consumer risk appetite, the Credit Cycle, and the Housing Bubble”, *Working Papers, Research Department, Federal Reserve Bank of Philadelphia*, No. 16-05.
7. Breeden, J.L. and L.C. Thomas (2016), “Solutions to Specification Errors in Stress Testing Models”, *Journal of the Operational Research Society*, June, p 830-840.
8. Breeden, J.L., L. Thomas, and J.W. McDonald III (2008), “Stress-testing retail loan portfolios with dual-time dynamics,” *The Journal of Risk Model Validation,* ***2****(2),* Summer*, pp. 43-62.*
9. Coffman, John Y., and Gary G. Chandler (1983), *Applications of performance scoring to accounts receivable management in consumer credit*. Credit Research center, Krannert Graduate School of Management, Purdue University.
10. Cox, D. R. (1972), “Regression models and life-tables”. *Journal of the Royal Statistical Society. Series B (Methodological),* **34**:2, pp.187.
11. Cox, D. R., & Oakes, D. O. (1984), *Analysis of survival data*. London: Chapman and Hall.
12. Financial Accounting Standards Board (2016), “Financial Instruments – Credit Losses (Topic 326): Measurement of Credit Losses on Financial Instruments”, *Financial Accounting Series*, FASB Accounting Standards Update, No. 2016-13, June.
13. Glenn, N.D. (2005), *Cohort Analysis*, 2nd ed. London: Sage.
14. Holford, T. R. (1983), The estimation of age, period and cohort effects for vital rates. Biometrics, **39**, pp. 311–324.
15. Judge, G., Griffiths, W. E., Hill, R. C., Ltkepohl, H., & Lee, T.-C. (1985), *The theory and practice of econometrics.* Wiley.

16. McPhail, J., L. McPhail, and B. Bazyluk, (2015), "Considerations for Determining the Allowance for Loan and Lease Losses", *The RMA Journal*, **97**:8, May, pp. 44-49.
17. Øksendal, B. (2000), *Stochastic Differential Equations: An Introduction with Applications, Fifth Edition*, Springer-Verlag: New York.
18. Schmid, V.J., and L. Held (2005) ,"Bayesian Age-Period-Cohort Modeling and Prediction — BAMP," *Journal of Statistical Software*, **21**:8, pp. 1–15.
19. Thomas, L. C. and J. Ho, and W. T. Scherer (2001), "Time will tell: behavioural scoring and the dynamics of consumer credit assessment", *IMA Journal of Management Mathematics*, **12**:1, p 89.