1-1-2017

What Can We Learn from Past Mistakes? Lessons from Data Mining the Fannie Mae Mortgage Portfolio

Stanislav Mamonov
Montclair State University, mamonovs@mail.montclair.edu

Raquel Benbunan-Fich
Baruch College

What Can We Learn from Past Mistakes? Lessons from Data Mining the Fannie Mae Mortgage Portfolio

Authors Stanislav Mamonov and Raquel Benbunan-Fich

Abstract Fannie Mae has been widely criticized for its role in the recent financial crisis, yet no detailed analysis of the systematic patterns of the mortgage defaults that occurred has been published. To address this knowledge gap, we perform data mining on the Fannie Mae mortgage portfolio of the fourth quarter of 2007, which includes 340,537 mortgages with a total principal value of $69.8 billion. This portfolio had the highest delinquency rate in the agency's history: 19.4% versus the historical average of 1.7%. We find that although a number of information variables that were available at the time of mortgage acquisition are correlated with the subsequent delinquencies, building an accurate model proves challenging. Identification of the majority of delinquencies in the historical data comes at a cost of low precision.

The financial crisis of 2007–2009 is considered the worst since the Great Depression of the 1930s (Financial Crisis Inquiry Commission, 2011). The crisis was precipitated by the rapid decline in housing prices in the United States. The decline triggered a complex web of events leading to the insolvency of a number of financial institutions and consequent freezing in the credit markets, which had broad effects across the economy (Financial Crisis Inquiry Commission, 2011). The U.S. GDP contracted 0.3% in 2008 and another 3.1% in 2009 (Young, 2013). The financial crisis affected many people. Nearly $11 trillion in household wealth vanished, and four million families lost their homes to foreclosure (Young, 2013). Unemployment reached 10% in the fall of 2009 (Bureau of Labor Statistics, 2015) and the effects of the crisis persist eight years later (Andriotis, 2015).

The rapid decline in housing prices starting in 2007 has been attributed to the period of irrational exuberance in the preceding years that was fueled by easy credit available for home financing (Shiller, 2015). Approximately 70% of real estate purchases in the U.S. are financed (Financial Crisis Inquiry Commission, 2011), meaning that the buyers borrow at least a part of the home value to facilitate the purchase. Government-sponsored enterprises (GSEs), Fannie Mae and Freddie Mac, were established in 1938 and 1970 respectively, to make it easier for individual homebuyers to afford a home (Peterson, 2008). The GSEs buy mortgages from banks and financial intermediaries, offering liquidity in the mortgage markets and making it easier for individual homebuyers to acquire financing.

The GSEs back nearly 60% of all individual real estate mortgages in the U.S. (Kan and Robotti, 2007). The agencies securitize mortgages for sale to investors and they also hold a substantial portfolio of mortgages on their balance sheets. Fannie Mae is the larger of the two agencies. In 2007, the Fannie Mae mortgage portfolio was valued at $403 billion, while the Freddie Mac portfolio was valued at $75 billion (Fannie Mae, 2008; Freddie Mac, 2008). Both agencies were highly leveraged and the decline in the real estate values associated with the financial crisis triggered a wave of mortgage defaults that effectively bankrupted both agencies, leading the Federal Housing Finance Agency to place both in conservatorship in 2008 (Financial Crisis Inquiry Commission, 2011).

The GSEs have been frequently criticized for loose underwriting standards in the period preceding the crisis (Wallison and Calomiris, 2009). However, to the best of our knowledge, no systematic analysis of the agencies' mortgage portfolios has been published to substantiate the criticism, and, even more importantly, to extract lessons for the future. We take the first steps in this task here. We review the Fannie Mae prime 30-year fixed-rate mortgage (FRM) portfolio delinquencies in the 2000–2014 period and we identify the fourth quarter of 2007 as the mortgage portfolio with the highest rate of severe delinquencies. We perform extensive data mining on this portfolio to understand the extent to which data mining techniques can be used to build predictive models based on the identification of systematic patterns and salient predictors. The answers to these questions shed light on the systematic relations between the information variables that were available prior to mortgage origination and mortgage delinquencies during the financial crisis to inform practice and policy decisions. In the sections that follow, we provide an overview of the studies on mortgage delinquencies associated with the recent financial crisis, as well as a summary of prior work on applying data mining techniques to predict credit defaults. We describe the dataset in our study and present the data mining results. We conclude with a discussion of our findings and their implications for practice and policy.

Background

Government estimates for the fourth quarter of 2015 show that there is approximately $13.795 trillion in outstanding mortgage obligations in the U.S. (Federal Reserve, 2016). Mortgage lending is an important area of practice and the recent financial crisis stimulated new research aimed at understanding the causes of the mortgage defaults (Demyanyk and Van Hemert, 2011), as well as the evaluation of the effectiveness of the government programs aimed at alleviating borrower hardship brought about by the financial crisis (Schmeiser and Gross, 2016). The research on mortgage defaults is commonly grounded in the competing hazards model developed by Deng (1997) and Clapp, Deng, and An (2006) who apply econometric models to investigate the individual and structural factors that affect mortgage default decisions. Much of the published research in this stream relies on the information that was available after the mortgages had been issued (ex post stage). The research done by Smith (2011) exemplifies the use of the ex post data for analysis. Using loan-level and credit data to evaluate mortgage performance after origination, the author shows that declining credit scores are associated with the higher probability of a default, whereas increasing credit scores are associated with refinancing.

There has been much less recent work examining the effects of the salient factors that are available prior to mortgage origination (ex ante stage). This is likely due to the transition of many financial institutions to automated underwriting systems that in effect codified existing practices (Lacour-Little, Park, and Green, 2012). For example, the maximum loan-to-value (LTV) ratio of 80% is required by GSEs (Consumer Financial Protection Bureau, 2013), and it has become a widely accepted industry standard for qualified mortgages. Borrowers are generally required to obtain mortgage insurance if they borrow more than 80% of the value of a home, thus significantly raising the cost of obtaining a non-qualified mortgage. Similarly, the industry has established norms for the qualified mortgage borrower credit score, LTV, and debt-to-income (DTI) as other critical factors (Consumer Financial Protection Bureau, 2014). To the best of our knowledge, there has not been a published systematic examination of the critical values in these factors that may influence mortgage defaults, particularly in the environment of falling real estate prices and increasing unemployment. Further, there is some disagreement on the critical role of these factors in lending decisions. For example, Archer, Elmer, Harrison, and Ling (2002) show that LTV had little predictive value in multifamily properties.

In this study, we examine the data that were available to Fannie Mae prior to mortgage origination (ex ante) and we apply data mining techniques to explore the systematic relations that were present at the mortgage origination stage that may yield clues to mortgage risks. We also examine the effects of specific critical values of the borrower credit score, LTV, and DTI on mortgage defaults in the Fannie Mae portfolio.

Research on the Mortgage Defaults Associated with the Recent Financial Crisis

The competing hazards model, developed by Deng (1997), Ciochetti, Deng, Gao, and Yao (2002), Clapp, Deng, and An (2006), An and Qi (2012), and Jiang, Nelson, and Vytlacil (2014), is the dominant theoretical perspective in the recent ex post mortgage default research. In the competing hazards model, lenders face two interdependent risks. The first risk is that the borrower will repay the mortgage ahead of term, thus precluding the lender from earning interest over the full term of the mortgage. The second risk is that the borrower will default on the mortgage. Clearly, the mortgage default risk poses a greater threat because it affects both the unpaid principal and the future interest payments. The two risks are interdependent, because an early mortgage prepayment eliminates the default risk and a mortgage default markedly reduces the likelihood of an early prepayment.

In the wake of the recent financial crisis, there have been investigations into the factors that affect mortgage default decisions that revealed some unexpected findings. Initial studies suggested that the decline in housing prices produced negative equity (outstanding mortgage balance being higher than the value of a home) for many borrowers. The negative equity was proposed as the motive underlying the rising rate of defaults (Bajari, Chu, and Park, 2008; Foote, Gerardi, and Willen, 2008). Subsequent research on the mortgage defaults in 2008 showed that negative equity was not an immediate trigger for a mortgage default, as may be expected from a completely rational real estate investor, but rather most homeowners with negative equity did not default until the negative equity reached 40% of the value of the home (Campbell and Cocco, 2011). Elul et al. (2010) further elucidated the relation between negative equity and the defaults by showing that liquidity shocks (loss of income) play a greater role in explaining mortgage defaults than negative equity. To add a further nuance to the complexity of the individual decisions underlying mortgage defaults, a survey of mortgage borrowers showed that individual numerical ability is negatively correlated with mortgage defaults after controlling for general cognitive ability, as well as demographic and financial variables (Gerardi, Goette, and Meier, 2010). While the studies focusing on the ex post default decisions offer insights on the underlying causes of mortgage defaults, the ex post data (e.g., a borrower's employment prospects during an economic downturn) are difficult to gauge accurately at the time of mortgage origination and therefore, these results are difficult to translate into underwriting decisions.

Data Mining Studies of Credit Defaults

Data mining, also often called "knowledge discovery in databases" (KDD) refers to algorithmic discovery of patterns in data (Fayyad, Piatetsky-Shapiro, and Smyth, 1996). We should note that data mining is different from the commonly employed econometric models that are concerned with parameter estimation in the context of specified models. Data mining techniques do not specify a model a priori, but rather focus on evaluation of how different types of data mining models can capture patterns of covariation in the data. Published data mining studies of mortgage defaults have been primarily done using datasets originating from outside the U.S. using different modeling techniques (support-vector machine, artificial neural network, decision tree, and random forest). A seminal study that evaluated the efficacy of different modeling techniques using eight credit scoring datasets from the United Kingdom and Benelux suggest that support vector machines and artificial neural network algorithms could deliver the best results (Baesens et al., 2003). A study of mortgage defaults in Israel found that the decision tree algorithm offered the best accuracy in predicting defaults (Feldman and Gross, 2005), and a study of a synthetic German credit dataset based on real-world data showed that random forest algorithm outperforms other techniques in predicting loan defaults (Ghatasheh, 2014). In summary, data mining studies with international credit datasets did not produce conclusive findings regarding the best way to model credit defaults. This has been confirmed in recent work that showed that different algorithms offer better performance across different international credit-related datasets (Zurada, Kunene, and Guan, 2014).

Method

Data Source

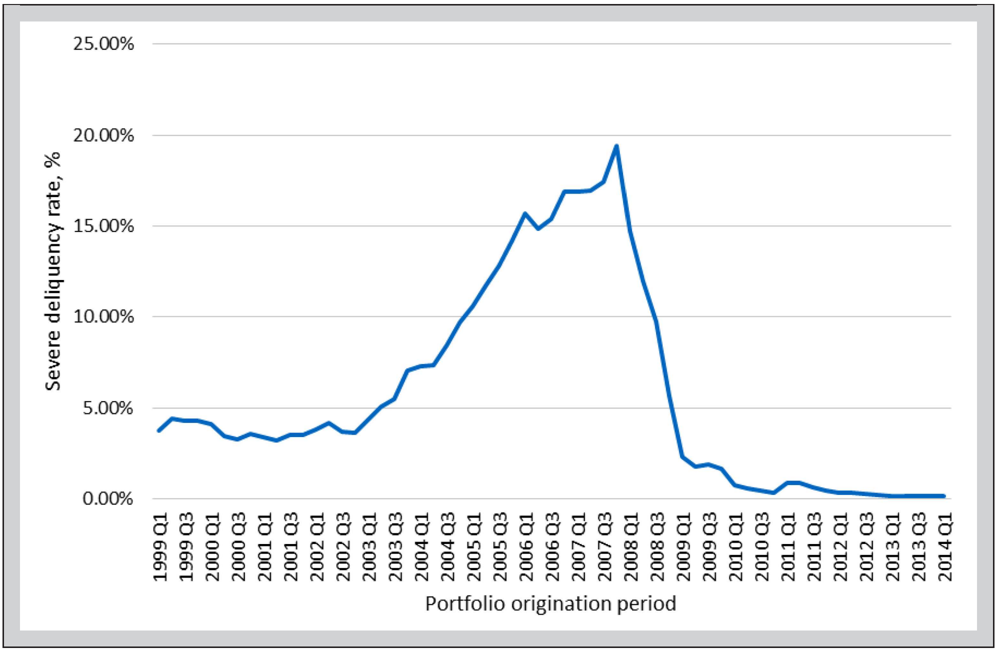
Following the financial crisis, GSEs are required to make their mortgage origination and mortgage performance data public. We obtained the Fannie Mae mortgage origination and mortgage performance data covering the period between the first quarter of 2000 and the first quarter of 2014 directly from the agency. The mortgage origination dataset contains information that was available to the agency at the time of mortgage acquisition. These data include individual borrower characteristics (e.g., personal credit score), as well as information about the property (number of units) and the financial details of the transaction (e.g., LTV ratio). The complete data dictionary is provided in the Appendix. The full dataset includes 21.7 million FRMs with the combined principal value of $4.186 trillion acquired by Fannie Mae between January 2000 and March 2014.

The mortgage performance dataset contains information about how the specific loans performed over time after acquisition by Fannie Mae on a monthly basis. The dataset contains over 917 million records pertaining to 21.7 million individual mortgages. Each record in the mortgage performance dataset contains the Loan Identifier field that is related to the Loan Identifiers specified in the mortgage origination dataset. This correspondence allowed us to relate the mortgage origination data to the mortgage performance data.

Industry practice shows that mortgage payers who fall behind by three months nearly invariantly end up in default on the mortgage obligation (Sun, 2013). The three-month period of delinquency is often referred to as "technical default" in the banking industry (Quercia and Stegman, 1992). However, to avoid confusion with the actual mortgage default that requires the transfer of legal rights to the property and is often delayed in relation to the technical default (Allen, Peristiani, and Tang, 2013), we refer to a three-month delinquency as a "severe delinquency."

To develop the dataset for our analysis, we combined the information containing the predictor variables from the mortgage origination dataset with the subsequent delinquency status of the individual mortgages from the mortgage performance

Exhibit 1 | Severe Delinquency Rates in the Fannie Mae Portfolios in 2000:Q1–2014:Q1



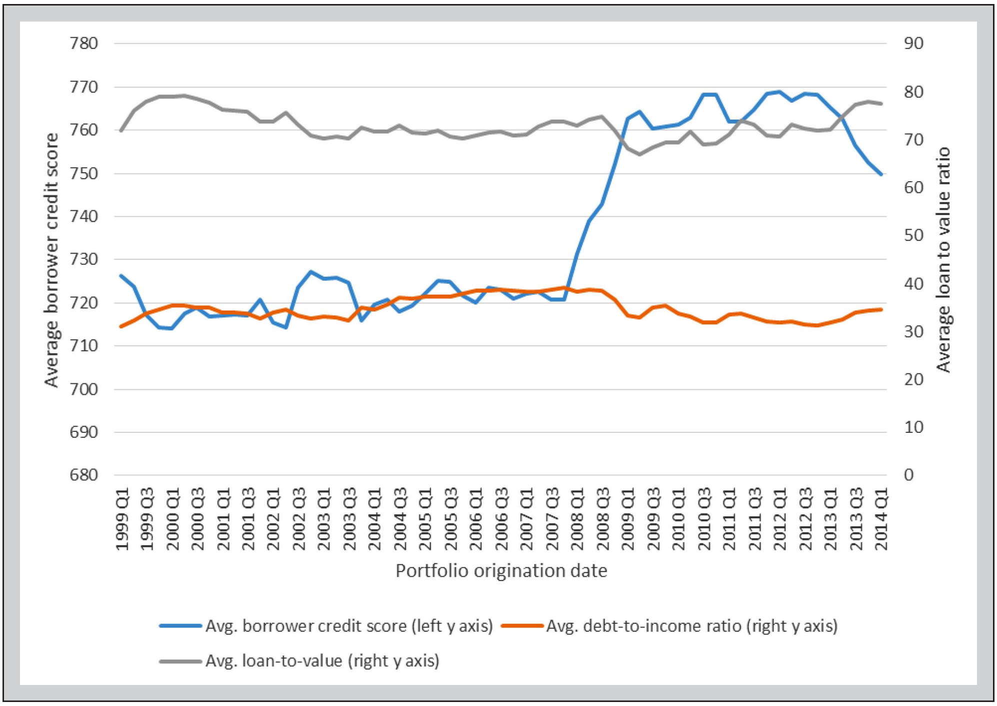
dataset. We created a binary dependent variable, Severe Delinquency, which we assigned the value of 1 if a loan became delinquent for three or more months and 0 otherwise.

Exploratory Analysis

In the first step of our analysis, we examined the historical delinquency rates for mortgages acquired by Fannie Mae over the period from 2000:Q1 to 2014:1. In Exhibit 1, we summarize the severe delinquency rates for the FRMs acquired by Fannie Mae over this period of time. As can be seen, the delinquency rate rose dramatically for mortgages acquired by the agency through the end of 2007. The portfolio of mortgages acquired by Fannie Mae in 2007:Q4 had the highest default rate over the history of the agency at 19.4%. The historical mortgage default rates in the Fannie Mae portfolio averaged 1.7% (Peterson, 2009). The Fannie Mae 2007:Q4 prime mortgage portfolio includes 340,537 mortgages, with a total principal value of more than $69.8 billion.

In the next step, we examined the historical variability of the key factors known to affect mortgage defaults: borrower credit scores, LTV, and DTI (Demyanyk and Van Hemert, 2011). The plot of the results for mortgages acquired by Fannie Mae in 2000–2014 presented in Exhibit 2 does not reveal any drastic changes in the average borrower characteristics (credit scores) or the loan characteristics (LTV

Exhibit 2 | Average Borrower Credit Score, LTV, and DTI in the Fannie Mae Portfolios in 2000:Q1–2014:Q1



or DTI) over the period preceding 2007:Q4. There is a significant rise in the average borrower credit score following the financial crisis, reflecting credit tightening that occurred in its aftermath (Shenn, 2012), but no obvious deterioration in the borrower credit scores, financial leverage (DTI) or increasing amount of borrowing vis-a-vis the value of the properties (LTV) are evident in the period prior to 2007:Q4.

The exploratory analysis did not produce any obvious insights into the potential causes of the significant rise in the delinquency rates in the Fannie Mae portfolio of mortgages in 2007–2008. This raises the question of whether there are systematic patterns of delinquencies in the Fannie Mae portfolio that can shed light on the underlying causes of delinquencies and help prevent similar events in the future. To address this question, we performed data mining on the dataset of mortgages acquired by Fannie Mae in 2007:Q4, seeking to build predictive models able to capture patterns in the mortgage delinquencies that occurred. Our rationale for choosing to focus on this dataset stems from the fact that this mortgage portfolio had highest delinquency rate in the agency history at 19.4%, versus the historical average of 1.7%. This dataset bears witness to the course of mortgage delinquencies that followed the financial crisis. Analysis of this dataset may yield

Exhibit 3 | Summary Statistics

Variable	Summary Statistics
Original interest rate	Mean = 6.51%, Std. dev. = 0.37%
Original balance	Mean = $205,327, Std. dev. = $100,687
Loan-to-value (LTV) ratio	Mean = 73.73, Std. dev. = 15.73%
Combined loan-to-value (CLTV) ratio	Mean = 75.37, Std. dev. = 16.07%
Number of borrowers	Mean = 1.52, Std. dev. = 0.52
Debt-to-income (DTI) ratio	Mean = 39.30, Std. dev. = 12.28
Borrower credit score	Mean = 719, Std. dev. = 61
Co-borrower credit score	Mean = 727, Std. dev. = 60
First time borrower	Yes = 11.1%, No = 88.8%, Unknown = 0.1%
Loan purpose	Purchase = 41%, Cash-out refinance = 39.4%, Refinance = 19.6%
Property type	Single-family homes = 73.4%, Planned unit development = 15.8%, Condo = 9.5%, Multi-family homes = 0.7%, Co-op = 0.6%
Number of units	1 = 96.7%, 2 or more = 3.3%
Occupancy	Principal = 85.6%, Investment = 9.6%, Second home = 4.8%

insights to the value of information variables available prior to mortgage origination that can affect subsequent mortgage delinquencies.

Close inspection of the dataset revealed that the data provided by Fannie Mae is generally of good quality with few aberrant records. We removed 530 records that were missing the primary borrower credit score and 99 records that had $0 original balance (these records were a subset of the records missing the credit score). After cleaning, the dataset contains 340,007 records, with a combined principal value at origination of $69.8 billion. Exhibit 3 below provides the summary statistics for the individual variables in the data.

Prediction Models

Loan delinquency prediction is a binary classification problem. Prior research has shown that different data mining algorithms perform better on different loan datasets (Zurada, Kunene, and Guan, 2014). We evaluated six data mining algorithms for their ability to predict loan delinquencies in our sample: logistic regression, decision tree, random forest, boosted trees, support-vector machines, and artificial neural networks. We briefly discuss each of the modeling techniques, as well as their merits and weaknesses. In the discussion of merits and weaknesses, we specifically focus on the interpretability of individual models, because mortgage delinquency prediction may expose the agencies to legal challenges requiring the agencies to justify their decision to accept or reject a specific mortgage. For this reason, an ideal model would provide full transparency into the mortgage acceptance/rejection decision.

Logistic regression is a generalization of the linear regression models. This modeling technique relates the predictor variables to the log of odds of an event, and it estimates the parameters using the maximum likelihood approach for the log of odds function:

$$Log(Odds) = \beta_0 + \beta_1 x_i + \beta_2 x_2 + \beta_i x_i,$$

where x_i [$i = 1, 2, \ldots, n$] are variables that influence the odds of the outcome of interest.

In a binary classification problem with a balanced sample, odds greater than 50% would mean that the event will occur, and odds below 50% would mean that the event would not occur. Logistic regression is a popular modeling technique frequently used in practice (Hosmer and Lemeshow, 2004), although the method is not without limitations. For example, a linear relation is assumed between the predictors and the log of odds of an event occurrence in a linear regression. Logistic regression provides visibility into the significance of individual predictors in the model and the sign (positive or negative) associated with each; however, the specific correlation coefficients may be difficult to interpret. Further, the logistic regression model is sensitive to missing values. Data imputation is required to retain cases with missing values and this can lead to biased parameter estimates, particularly where the missing values are non-random (Allison, 2000).

Decision trees is a classification algorithm that recursively separates observations in branches to build a tree for the purpose of improving prediction accuracy (Safavian and Landgrebe, 1991). For classification problems, the branching points are based on the improvement in one of the commonly used information gain metrics (Entropy or Gini index), which capture the improvement in homogeneity of each subset of data after the split. The branching points identify the variables and the corresponding thresholds that are used for the data split. The decision tree models are transparent and easy to interpret; however, the decision tree algorithm is greedy and therefore it may not capture the optimal global partitioning of the data (Safavian and Landgrebe, 1991). While the decision tree algorithm is a powerful modeling technique, it has a known weakness in potentially over-fitting the training data (Bradford et al., 1998). A number of decision tree-based modeling techniques have been developed that leverage the decision tree ability to capture non-linear relations in the data and also safeguard against over-fitting.

Random forest is an example of an ensemble modeling technique that combines the predictions of multiple decision trees to achieve better overall performance (Breiman, 2001). The random forest algorithm builds multiple tree models by randomly selecting a subset of predictor variables and a subset of data to build each tree. The algorithm sets aside an ''out-of-bag'' subsample of data, continually evaluates the incremental improvement in the prediction accuracy with the addition of each new tree, and it only retains the trees that to the overall model accuracy.

Boosted trees is another tree-based ensemble modeling technique (Bauer and Kohavi, 1999). Similarly to the random forest models, the boosted trees algorithm involves the construction of multiple tree models and aggregating the predictions across the collection of models. The distinction of the boosted trees approach to modeling is in improving the prediction accuracy by increasing the weights of misclassified cases in each modeling round. By focusing on the misclassified cases, the boosted trees algorithm can develop better overall results. By virtue of being ensemble techniques, both the random forest and boosted trees models offer limited visibility into how individual predictors affect the dependent variable, but the models can be used to identify the relative importance of the individual predictors to the overall model quality.

In addition to the models discussed above, which provide at least a degree of transparency into the effects of individual factors, we also include two ''black box'' modeling techniques in our analysis: support vector machine (SVM) and artificial neural networks (ANN). The SVM modeling technique applies mathematical (kernel) functions to transform the input feature space to identify a boundary that can help separate the two classes of outcomes for a binary variable (Amari and Wu, 1999). SVMs can utilize different kernel functions. Following the evaluation of different kernels, we found that the Laplacian transform (Qi, Tian, and Shi, 2012) produced the best results for the SVM family of models with our dataset and this is the kernel function for which we report the SVM results. The ANN is another modeling algorithm that we utilized. ANNs are an advanced modeling technique that evolved from research aiming to model the function of biological neural networks (Yegnanarayana, 2009). ANNs are typically comprised of several layers of interconnected nodes. The input layer nodes correspond to the individual predictor variables. The input nodes are connected to the inner layer nodes, which can be programmed to perform different types of non-linear transformations (e.g., a logistic function), and which, in a binary classification problem, ultimately connect to a single output node. The parameters affecting the individual connections between the nodes in the neural networks are ''learned'' through training by utilizing a back-propagation function, which captures the errors on the output node and back-propagates the parameter adjustment throughout the network to achieve better fit over the training rounds.

Model Performance Evaluation

The performance of the binary classification algorithms is commonly assessed by splitting the data, using a part of the data (training set) to build the models and

Exhibit 4 | Classification Matrix

	Predicted	
	False	True
Actual		
False	True negatives (TN)	False positives (FP)
True	False negatives (FN)	True positives (TP)

then assessing the model performance on the remaining data (test set) using the classification matrix (Exhibit 4) and the derived metrics. In this study, we relied on three model performance measures: true positive rate, positive predictive value rate, and accuracy.

$$\text{True positive rate} = TP/(TP + FN)$$

$$\text{Positive predictive value} = TP/(TP + FP)$$

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN)$$

where TP denotes true positives, TN denotes true negatives, FP denotes false positives, and FN denotes false negatives.

The true positive rate reflects the model's ability to correctly identify severe delinquencies that occurred vis-à-vis false negative errors, which indicate mortgages that are predicted to not fall into delinquency, but did. A model with the higher positive predictive value will make fewer false positive errors. The positive predictive value reflects the model's ability to correctly predict true positives, which indicate mortgages that become delinquent, vis-à-vis false positives. The false positive signal from a model would imply that a mortgage is likely to become delinquent and such errors would likely lead to denial of credit to misclassified borrowers.

It is important to note that the true positive rate of the individual models is a critical measure of the model's performance for predicting loan delinquencies. A false negative, a mortgage delinquency that is not predicted accurately at origination, exposes the underwriter to the potential loss of a part of the principal. The average value of a loan in the 2007:Q4 portfolio was approximately $205,000. Prior research on actual losses by mortgage underwriters in case of a default (loss given default, LGD) suggests that the LGD can reach 50% on residential mortgages (Park and Bang, 2014). Most of the severe delinquencies in our dataset occurred within 18 months of the mortgage origination when less than 5% of the principal had been repaid. A conservative estimate of 25% LGD would imply that every false negative would carry a cost of at least $50,000 without accounting for the costs associated with the asset recovery. In recognition of the true positive rate as the key metric for the evaluation of model performance, we optimized our models for the maximum true positive rate, while setting the positive predictive value threshold to 30%.

We followed the recommended practice of k-fold cross-validation (Breiman, Friedman, Olshen, and Stone, 1984) for the evaluation of the individual algorithm performance. The k-fold cross-validation consists of three steps:

1. Dividing the dataset into k disjoint subsets. Subsets are stratified for the dependent variable to reduce model accuracy estimation bias (Kohavi, 1995).
2. Training the algorithm on k = 1 subsets while withholding one of the subsets for model performance evaluation. The process is repeated withholding each of the subsets.
3. The cross-validation model performance results are averaged to produce an estimate of the classifier accuracy.

Research suggests that 10-fold cross-validation is generally sufficient to establish model performance estimate (Breiman, Friedman, Olshen, and Stone, 1984; Kohavi, 1995).

We used R (64-bit, version 3.1.3) software to build and evaluate the data mining models (Anon, 2015). R includes an implementation of the general linear models (glm) in the default distribution. We used this implementation for the logistic model in our analysis. We used the following packages to build the respective models: rpart (decision tree), randomForest, ada (boosted trees), kernlab (SVM), and nnet (neural networks).

Results

All models had difficulty with the accurate prediction of mortgage delinquencies. The ANN neural network algorithm showed the best results, accurately predicting 83.4% of delinquencies on average. The random forest model performed the worst, accurately predicting just 58.9% of the delinquencies. The results of 10-fold cross-validation for each of the modeling techniques are given in Exhibit 5.

It is also important to note that while we optimized the models for the true positive rate, we sacrificed the positive predictive value, a.k.a. the precision of the models. The average precision of the ANN is only 36.7%. This means that roughly two out of three predicted severe delinquencies will be false alarms. These (false positive) errors would imply opportunity costs (missed opportunity to earn interest) and they would also potentially affect availability and cost of credit to the population of misclassified borrowers.

Exhibit 5 | Model Performance

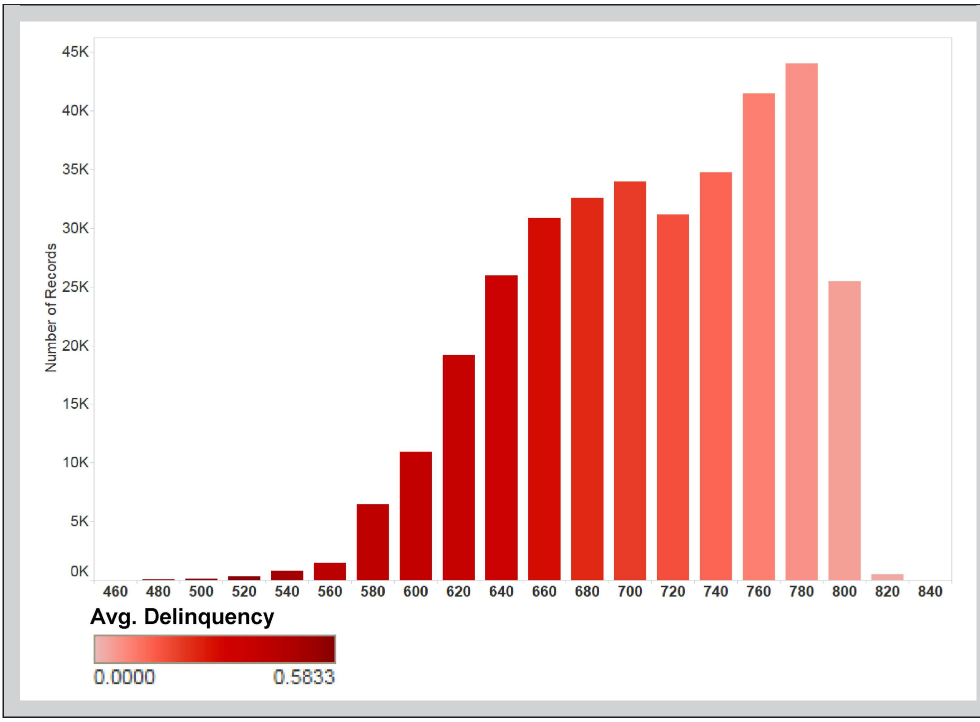
	True Positive Rate	Positive Predictive Value	Accuracy
Logistic regression	82.7% ± 0.8%	36.9% ± 2.4%	59.9% ± 0.6%
Decision tree	70.3% ± 2.7%	32.9% ± 2.9%	66.5% ± 0.7%
Random forest	58.9% ± 0.5%	30.9% ± 0.9%	66.5% ± 0.1%
Boosted trees	64.6% ± 1.5%	42.8% ± 2.0%	69.2% ± 0.7%
SVM	82.8% ± 3.2%	36.5% ± 2.6%	59.3% ± 0.5%
Neural network	83.4% ± 2.0%	36.7% ± 2.9%	59.4% + 0.5%

Exhibit 6 | Feature Importance for the Model's True Positive Rate

Neural Network		SVM		Logistic Regression	
Feature	Score	Feature	Score	Feature	Score
Co-Borrower Credit Score	0.160	Credit Score	0.078	Co-Borrower Credit Score	0.106
Credit Score	0.117	Loan Purpose	0.034	Credit Score	0.073
DTI	0.063	DTI	0.029	CLTV	0.063
LTV	0.043	LTV	0.028	Mortgage Insurance	0.041
Original Balance	0.032	Seller	0.025	DTI	0.036
Seller	0.024	CLTV	0.025	LTV	0.027
Original IR	0.021	Num Borrowers	0.022	Original Balance	0.014
Mortgage Insurance	0.009	Channel	0.016	Original IR	0.012
Property Type	0.008	Original IR	0.016	Seller	0.005
CLTV	0.007	Original Balance	0.016	First Time Borrower	0.001

Although all the models performed poorly, it is still possible to gain insights on the influence of individual predictor variables on the model accuracy. In the next step, we examined the effects of individual variables on the accuracy of the models using the feature permutation-based method (Altmann, Toloşi, Sander, and Lengauer, 2010). This method relies on withholding individual predictors and iteratively examining the effect of withholding the information on the model positive predictive rate using the test data. The feature importance scores were estimated over several trials. The scores did not change significantly over the trials and the representative scores are provided because the focus is on the relative feature importance as opposed to the specific scores associated with the individual features. Exhibit 6 provides the results.

Exhibit 7 | Number of Records and Default Propensity vs. Credit Score

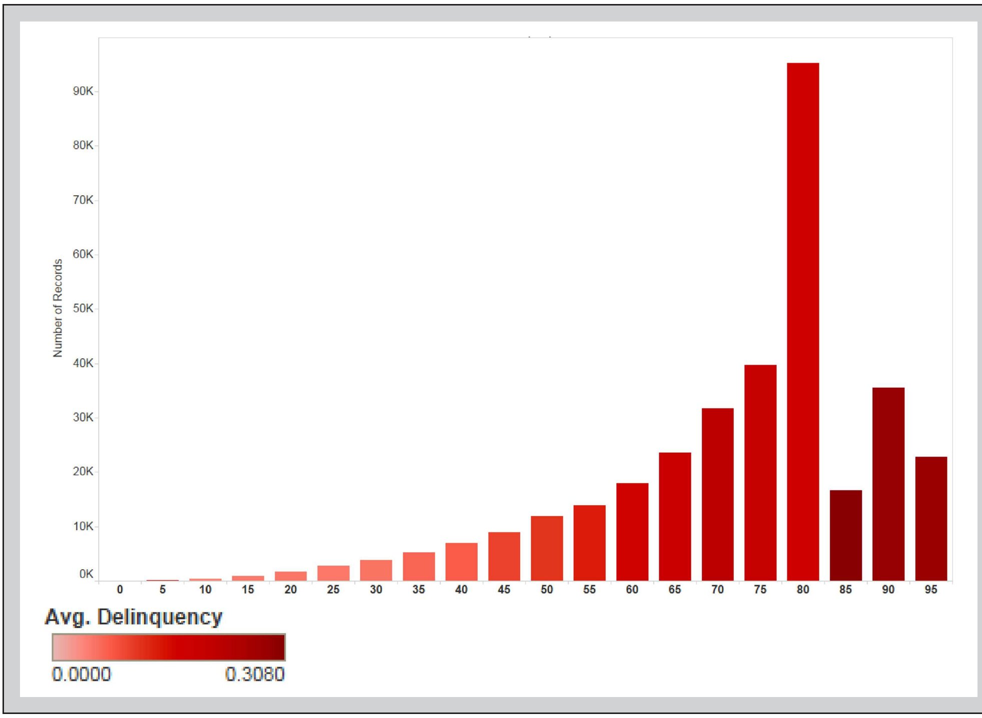


The borrower and co-borrower credit scores, LTV, CLTV, and DTI are well known predictors of mortgage defaults (Demyanyk and Van Hemert, 2011). To further explore the relation between these variables and subsequent delinquencies, we binned the records and visualized the number of records corresponding to each bin (the height of the bars in Exhibits 6–9) and the default propensity. Exhibits 10–12 reflect the data underlying the visualizations.

The visualizations reveal that the relations between the borrower credit score, LTV, and DTI are not linear. For example, the LTV visualization reveals that the default rate generally rises with the increasing LTV; however, the mortgages with LTVs of 75%–80% actually had lower delinquency rates (16.88%) versus mortgages with LTVs of 71%–75% (20.13%). The difference is significant at $p < 0.001$ (Z score = 13.996). The default rates rise dramatically for the mortgages with LTVs above 80% (30.8%). This difference is also statistically significant (Z = 42.2, $p < 0.001$).

To identify the critical values of these factors in influencing mortgage delinquencies in the Fannie Mae dataset, we constructed a decision tree model focusing on these variables. The resultant decision tree is shown in Exhibit 13.

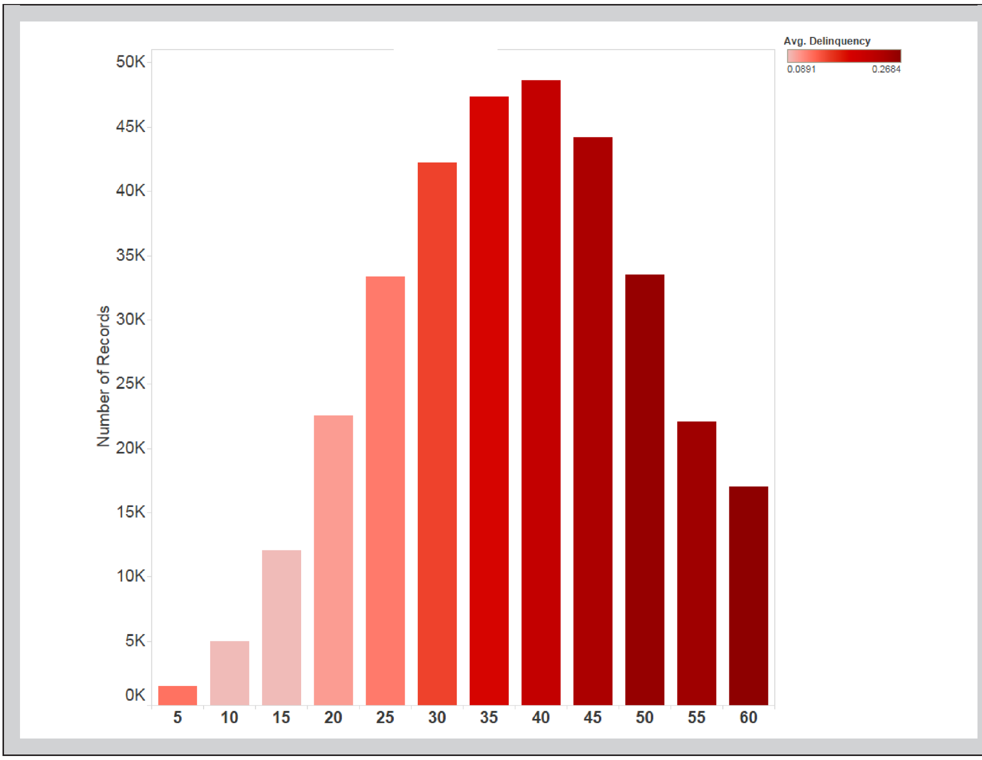
Exhibit 8 | Number of Records and Default Propensity vs. LTV



The decision rules emergent from the CART decision tree algorithm reveal that the credit score is the key information factor that is predictive of severe delinquencies. While the overall dataset has a 19.4% severe delinquency rate, FRMs issued to borrowers with credit scores greater or equal to 704 had only 11% severe delinquency, compared to 32% severe delinquency rate for borrowers with credit scores below 704. The severe delinquency rate increases to 38% for borrowers with credit scores below 666. The decision tree also reveals the layering of risks. LTV reflects the borrowed amount in relation to the value of the property; for borrowers with credit scores between 666 and 704, LTV is the key factor that is correlated with severe delinquencies. Borrowers with credit scores between 666 and 704, who borrowed more than 84% of the value of the property, exhibited the severe delinquency rate of 34%. For borrowers with credit scores below 666, DTI becomes a significant predictor of mortgage delinquencies; 41% of borrowers with credit scores below 666 and DTI greater than 34 ended up in severe delinquency.

Exhibits 14 and 15 provide further evidence of risk layering between the individual borrower credit score, individual level of indebtedness reflected in DTI and the equity held in the property at origination (LTV).

Exhibit 9 | Number of Records and Default Propensity vs. DTI



Discussion

In this study, we examined whether data mining techniques can capture systematic patterns of mortgage defaults in the Fannie Mae FRM portfolio. We focused on the mortgages acquired by the agency in the fourth quarter of 2007. We found that the borrower credit score, LTV, and DTI were the most significant factors correlated with the subsequent severe delinquencies. The list of factors identified as significant predictors of mortgage delinquencies in our models suggests that mortgage underwriters, including Fannie Mae, are collecting information that can be useful in predicting future delinquencies. We found that certain threshold values of credit scores, LTV, and DTI are associated with significantly higher delinquency rates in the Fannie Mae portfolio from 2007:Q4. Borrower credit scores below 704 were a strong predictor of serious delinquency. About a third (32%) of borrowers with credit scores below 704 were in technical default on their mortgages in our dataset, compared with less than 11.5% of borrowers with credit scores above 704. The size of the mortgage in relation to the property value (LTV) was also a significant predictor, but the relation between LTV and severe

Exhibit 10 | Number of Records and Delinquency Rate by Credit Score

Credit Score	Avg. Delinquency	# of Records
460	58.3%	24
480	56.8%	37
500	58.0%	143
520	57.9%	330
540	51.2%	772
560	42.8%	1,440
580	41.7%	6,467
600	40.1%	10,935
620	38.9%	19,167
640	34.9%	25,952
660	29.7%	30,853
680	24.3%	32,528
700	20.5%	33,985
720	16.9%	31,168
740	13.1%	34,759
760	8.9%	41,448
780	5.8%	44,034
800	3.8%	25,441
820	2.2%	502
840	0.0%	4

delinquency is not linear. Borrowers putting down payments of less than 16% were much more likely to become severely delinquent on their obligations (29% severe delinquency rate) than borrowers with $75 < \text{LTV} \leq 80$, for whom the severe delinquency rate was 16.9%. However, the delinquency rate was also higher (20.1%) for borrowers with $70 < \text{LTV} \leq 75$, indicating non-linear relations between the predictors and the target. We also found evidence of layered risks. Overleveraged borrowers with relatively low credit scores (<666), for whom the combined monthly debt obligations exceeded 34% of their gross monthly incomes, were also much more likely to fall behind on their mortgage payments (41% severe delinquency rate). Spotty prior credit history, limited personal financial investment in the property, and excessive borrowing against income make perfect sense as predictors of mortgage delinquency. Following the financial crisis, Fannie Mae announced tighter credit requirements for qualified mortgages (Reuters, 2008; Shenn, 2012); the agency raised the minimum required credit score from 580 to 640 and raised the minimum required down payment to 20%.

Exhibit 11 | Number of Records and Delinquency Rate by LTV

LTV	Delinquency Rate	Number of Records
0	0	9
5	16.0%	169
10	4.0%	428
15	4.9%	956
20	5.2%	1,741
25	5.1%	2,786
30	5.5%	3,815
35	6.7%	5,182
40	7.5%	6,931
45	10.0%	8,919
50	11.5%	11,975
55	13.9%	13,850
60	16.4%	17,987
65	18.8%	23,560
70	22.6%	31,756
75	20.1%	39,748
80	16.9%	95,214
85	30.8%	16,650
90	28.7%	35,520
95	28.3%	22,811

The surprising finding from our analysis was that it is very challenging to build an accurate prediction model for mortgage delinquencies using the data from the Fannie Mae portfolio. Model optimization for the true positive rate was only possible at a significant cost to the precision of the model. Our best model, a neural network, had an 83.4% average true positive rate; however, the average precision of the model was only 36.7%. Operationalizing the model would carry significant opportunity costs for the agency because it would likely lead to refusal to underwrite a significant number of mortgages. The false positive errors would also imply a societal cost as the rejected applications would likely preclude an opportunity to own a home.

This brings up the next question: How can we improve the quality of the data mining models to predict severe mortgage delinquency? One possible reason for the challenges that we encountered in building the models could be information insufficiency. We may be missing key information that could help us build better

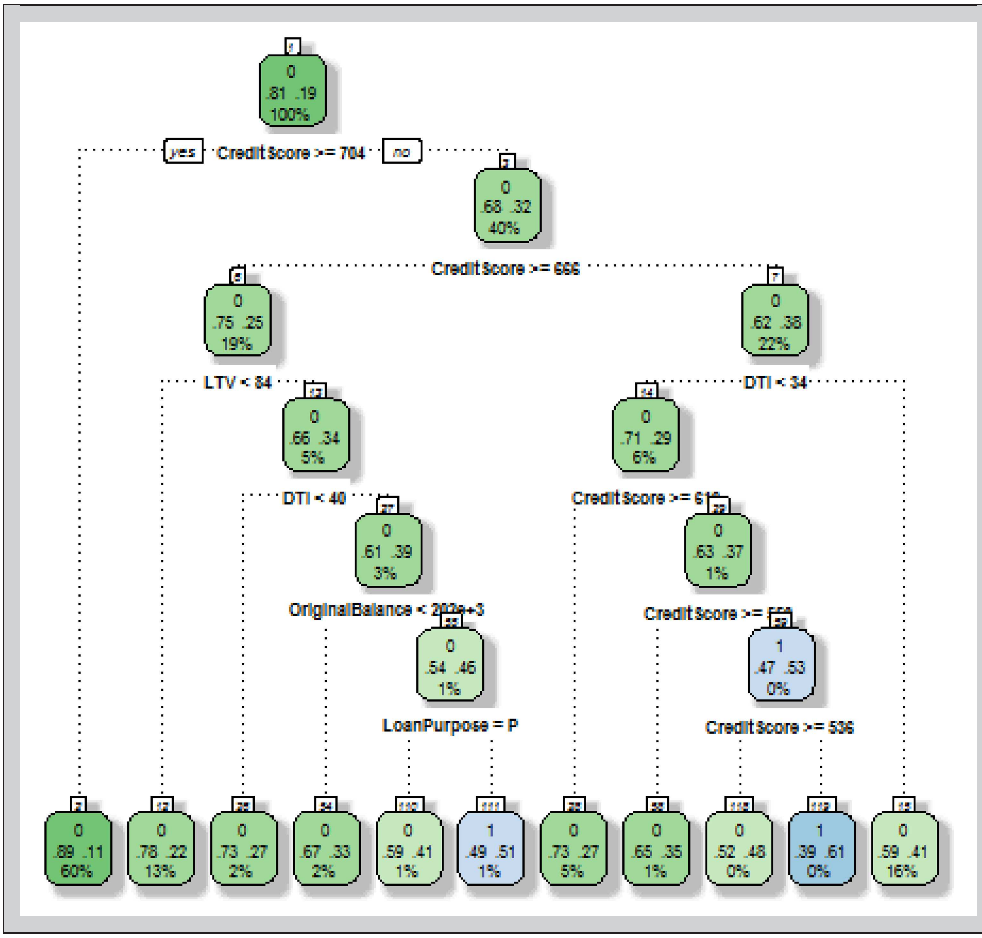
Exhibit 12 | Number of Records and Delinquency Rate by DTI

DTI	Delinquency Rate	Number of Records
5	12.5%	1,506
10	8.9%	4,977
15	8.9%	12,054
20	10.4%	22,501
25	12.1%	33,340
30	15.1%	42,256
35	18.7%	47,355
40	21.8%	48,611
45	24.2%	44,192
50	26.3%	33,464
55	25.5%	22,082
60	26.8%	17,017

models. Collection of additional information at the time of mortgage origination would offer a possible solution. Prior research offers some support for this proposal. Credit default analysis on a dataset from Israel, for example, identified the level of education and the type of professional employment as the key predictors of credit defaults (Feldman and Gross, 2005). Therefore, collection of additional information at the time of mortgage origination, including the education level and professional employment, may help improve the quality of the models.

An alternative and more likely explanation for the challenges that we encountered in building an accurate prediction model using the Fannie Mae mortgage portfolio dataset is that the financial crisis served as an exogenous cause of mortgage defaults. In this scenario, the information that was available at the time of mortgage origination simply would not be helpful in accurately predicting the consequences of a crisis for the portfolio. The exogenous cause explanation would imply that there was an external shock to the system that affected the base rate of mortgage delinquencies, as well as the nature of the deterministic and probabilistic relations among the data available at origination. Exogenous causes are often mentioned in discussions of macroeconomic models (e.g., models of unemployment) (Zivot and Andrews, 2002). The Great Depression and the oil crisis of the 1980s are classic examples of exogenous events that caused disruptions of linkages among macroeconomic factors and make it difficult to build accurate econometric models spanning these periods of history. The recent financial crisis had its origins in the rising defaults among the subprime borrowers that quickly spread to the prime mortgage borrowers and were amplified through

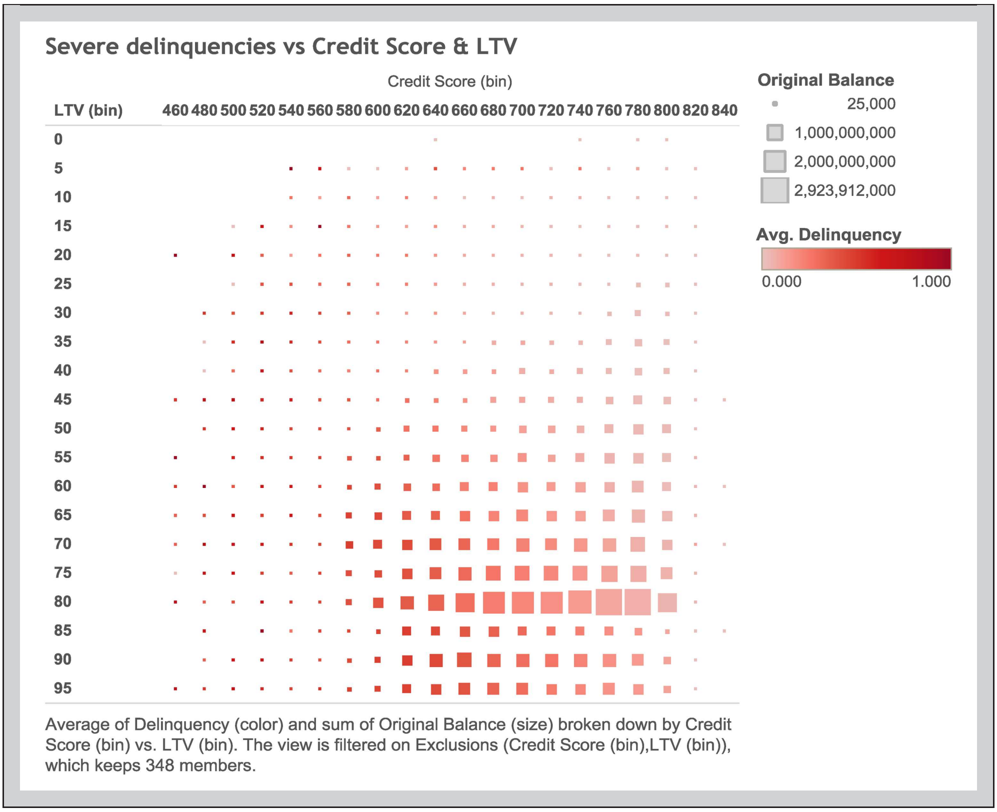
Exhibit 13 | Decision Tree Model Showing Critical Credit Score, LTV, and DTI Levels



the broader economic downturn (Financial Crisis Inquiry Commission, 2011). This exogenous shock reshaped the relations between the information used for borrower risk evaluation at the mortgage underwriting stage and subsequent defaults.

The recent financial crisis had a number of causes. The issuance of 5/1, 3/1, and 2/1 adjustable-rate mortgages (ARMs) and their securitization were among them (Financial Crisis Inquiry Commission, 2011). ARMs that carry a low introductory interest rate, which resets after the initial 2-, 3-, or 5-year period, gained in popularity in 2005–2006. Many of the ARMs were issued to subprime borrowers. The problems of subprime lending were also exacerbated by so-called "liar" loans (LaCour-Little and Yang, 2013) and corporate governance issues (Peni, Smith, and Vähämaa, 2013). As the interest rates on these mortgages began to reset in 2007, the mortgage payments for the borrowers grew drastically, triggering defaults (Mayer, Pence, and Sherlund, 2009). Although ARMs constituted a relatively

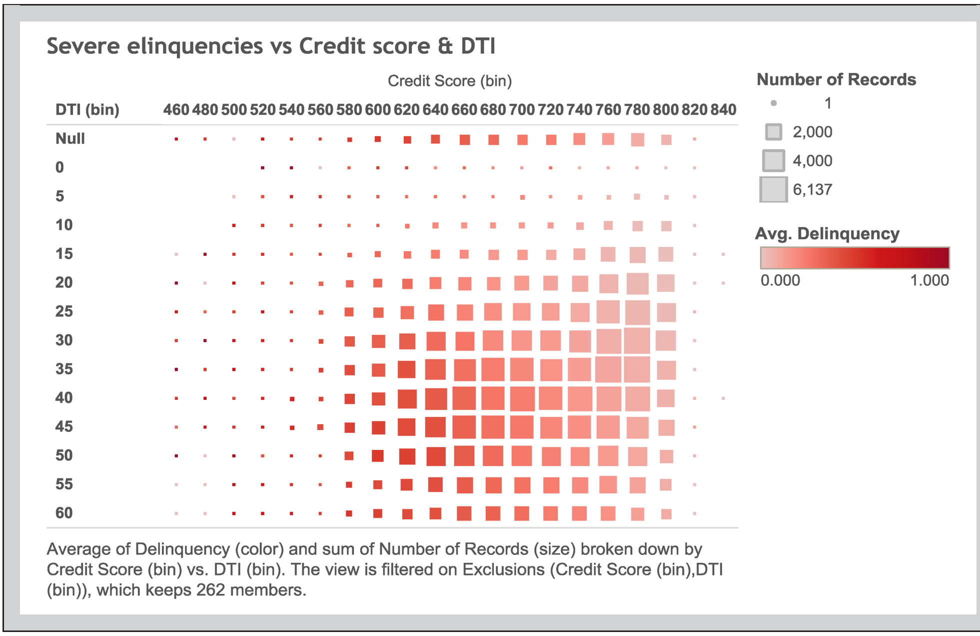
Exhibit 14 | Severe Delinquencies vs. Credit Score and LTV



small part of the overall mortgage market in 2007, the defaults on these mortgages produced a domino effect (Sherlund, 2010). As the properties bought with ARMs went into foreclosure they triggered rapid general declines in property values, as well as a series of events that affected all sectors of the economy, including prime mortgage borrowers. The economic downturn led to many people losing their jobs and the loss of steady income triggered many delinquencies on the traditional fixed-rate mortgages that were a part of the Fannie Mae portfolio.

The exogenous cause explanation for the failure of data mining techniques to accurately capture the patterns of defaults in the dataset imply that economic shocks will drastically increase mortgage default rates even among well-qualified borrowers. We examined the default rates among the best-qualified borrowers, those with credit scores above 760 for whom the historical delinquency rate across different types of credit (consumer loans, credit cards, mortgages, etc.) is less than 2% (Fair Issac Corporation, 2015). We find that the delinquency rate for this group was 5.5%–6.4% for the mortgages acquired by Fannie Mae in 2007. The 3X increase in the severe delinquency rate among the best qualified borrowers

Exhibit 15 | Severe Delinquencies vs. Credit Score and DTI



provides supporting evidence for the role of the financial crisis as an exogenous shock.

The current economic climate has largely pushed the concerns about the stability of the housing market to the back of everyone's mind and the GSEs have resumed some of the practices that contributed to the financial crisis. For example, the agencies now approve high LTV mortgages that require the borrowers to put just 3% down (Fannie Mae, 2015). The practical implication of the financial crisis being an exogenous shock is that even if Fannie Mae restricted mortgage purchases to the most qualified borrowers, the agency would have still faced bankruptcy. In this scenario, a significant reduction in the financial leverage of the agency would be necessary for the agency to weather the next financial crisis. The recent financial crisis brought the Dodd-Frank reform to the banking sector, effectively reducing financial average among the largest banks from 30:1 before the crisis to less than 10:1 after the reform (Acharya, Engle, and Richardson, 2012). A similar reform would be required to safeguard GSEs from bankruptcy going forward.

Conclusion

In this study, focusing on the data available prior to mortgage origination, we examined the predictive value of several data mining techniques using the Fannie

Mae mortgage dataset from the fourth quarter of 2007, which had the highest delinquency rate in the agency's history. Our data mining efforts revealed that the borrower credit score, and loan-to-value and debt-to-income ratios were the most important predictors of mortgage delinquencies. The ANN was the best performing model in our analysis. However, the ANN model identified the majority of severe delinquencies that occurred only at the expense of a high rate of false positives. The most likely reason for the predictive model shortcomings is that the financial crisis served as an exogenous shock, the effects of which cannot be accurately modeled using the data available at mortgage acquisition. This result suggests that Fannie Mae's current efforts to reduce future delinquencies by tightening mortgage qualification requirements may prove insufficient. Analysis of the Fannie Mae mortgage portfolio from 2007:Q4 shows that 16.7% of mortgages issued to borrowers with credit scores above 640 were severely delinquent in our dataset compared to the 1.7% historical delinquency rate. These results provide empirical support for the calls to reform the housing GSEs (Spahr and Sunderman, 2014).

Appendix

Data Dictionary for the Fannie Mae Mortgage Origination Dataset

Loan Identifier	Unique ID Assigned to Each Mortgage
Channel	The variable specifies the mortgage origination channel. The mortgages were either originated directly by retail banks, through a broker, or acquired from a different origination party after the mortgage was issued.
Seller name	The entity that delivered the mortgage loan to Fannie Mae. For legal reasons, we excluded this variable from our models.
Original interest rate	The original interest rate on a mortgage loan as identified in the original mortgage loan documents.
Original unpaid principal balance	The original amount of the mortgage loan as indicated by the mortgage documents. This is the amount of money being borrowed by the homebuyer to finance the purchase of the home.
Original loan term	The number of months in which regularly scheduled borrower payments are due.
Origination date	The date of the loan.
First payment date	The date of the first scheduled mortgage loan payment to be made by the borrower under the terms of the mortgage loan documents. The first payment date is typically 30–45 days after the mortgage origination date. This variable was not used in the models as it closely mirrors the Origination date.

Loan Identifier	Unique ID Assigned to Each Mortgage
Original loan-to-value (LTV)	A ratio calculated at the time of origination for a mortgage loan. The original LTV reflects the loan-to-value ratio of the loan amount secured by a mortgaged property on the origination date of the underlying mortgage loan. A higher LTV reflects that the homebuyer is borrowing a higher percentage of the property value.
Original combined loan-to-value (CLTV)	A ratio calculated at the time of origination for a mortgage loan. The CLTV reflects the loan-to-value ratio inclusive of all loans secured by a mortgaged property on the origination date. CLTV accounts for any secondary mortgages that the property owner may take out using the property as the collateral.
Number of borrowers	The number of individuals obligated to repay the mortgage loan.
Debt-to-income ratio	A ratio calculated at origination derived by dividing the borrower's total monthly obligations (including housing expense) by his or her stable monthly income.
Borrower credit score	A numeric value used by financial services industry to evaluate the quality of borrower credit. The score in the Fannie Mae portfolio is based on the "classic" FICO score developed by Fair Isaac Corporation.
First-time home buyer indicator	The indicator denotes whether or not a borrower or co-borrower qualifies as a first-time homebuyer. An individual is considered as a first-time homebuyer if he/she 1) is purchasing the property; 2) will reside in the property; 3) had no ownership interest in a residential property during three-year period preceding the date of the purchase of the property.
Loan purpose	An indicator that denotes if a mortgage is used for either property purchase, refinancing or refinancing with a cash-out option.
Property type	The field denotes whether the property is a cooperative share, condominium, planned urban development, single-family home or a manufactured home.
Number of units	The number of units comprising the related mortgaged property.
Occupancy status	The indicator denotes how the borrower used the mortgaged property at the origination date of the mortgage (principal residence, second home or investment property).
Property state	A two-letter abbreviation indicating the state within which the property securing the mortgage loan is located.
ZIP (3-digit)	The code designated by the U.S. Postal Service where the subject property is located.
Mortgage insurance percentage	The percentage of mortgage insurance coverage obtained for an insured conventional mortgage loan and used in the event of default to calculate the insurance benefit.
Co-borrower credit score	A numerical value used by the financial services industry to evaluate the quality of borrower credit. The score in the dataset refers to the "classic" FICO score developed by Fair Isaac Corporation.

References

Acharya, V., R. Engle, and M. Richardson. Capital Shortfall: A New Approach to Ranking and Regulating Systemic Risks. *The American Economic Review*, 2012, 102:3, 59–64.

Allen, L., S. Peristiani, and Y. Tang. Bank Delays in the Resolution of Delinquent Mortgages: The Problem of Limbo Loans. Fordham University Schools of Business Research Paper, 2013.

Allison, P.D. Multiple Imputation for Missing Data: A Cautionary Tale. *Sociological Methods & Research*, 2000, 28:3, 301–08.

Altmann, A., L. Toloşi, O. Sander, and T. Lengauer. Permutation Importance: A Corrected Feature Importance Measure. *Bioinformatics*, 2010, 26:10, 1340–47.

Amari, S. and S. Wu. Improving Support Vector Machine Classifiers by Modifying Kernel Functions. *Neural Networks*, 1999, 12:6, 783–89.

An, M. and Z. Qi. Competing Risks Models using Mortgage Duration Data under the Proportional Hazards Assumption. *Journal of Real Estate Research*, 2012, 34, 1–26.

Andriotis, A. Home-Equity Lines of Credit See Jump in Delinquencies. *The Wall Street Journal*, 2015.

Archer, W.R., P.J. Elmer, D.M. Harrison, and D.C. Ling. Determinants of Multifamily Mortgage Default. *Real Estate Economics*, 2002, 30:3, 445–73.

Baesens, B., T. Van Gestel, S. Viaene, M. Stepanova, J. Suykens and J. Vanthienen. Benchmarking State-of-the-Art Classification Algorithms for Credit Scoring. *Journal of the Operational Research Society*, 2003, 54:6, 627–35.

Bajari, P., C.S. Chu, and M. Park. An Empirical Model of Subprime Mortgage Default from 2000 to 2007. National Bureau of Economic Research, 2008.

Bauer, E. and R. Kohavi. An Empirical Comparison of Voting Classification Algorithms: Bagging, Boosting, and Variants. *Machine Learning*, 1999, 36:1–2, 105–39.

Bradford, J.P., C. Kunz, R. Kohavi, C. Brunk, and C.E. Brodley. Pruning Decision Trees with Misclassification Costs. In *Machine Learning: ECML-98*. Springer, 1998, 131–36.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone. Classification and Regression Trees. Monterey, CA: Wadsworth, Inc., 1984

Breiman, L. Random Forests. *Machine Learning*, 2001, 45:1, 5–32.

Bureau of Labor Statistics. Labor Force Statistics from the Current Population Survey. Databases, Tables & Calculations. Available at: http://data.bls.gov/timeseries/LNS14000000, 2015.

Campbell, J.Y. and J.F. Cocco. A Model of Mortgage Default. National Bureau of Economic Research, 2011.

Ciochetti, B., Y. Deng, B. Gao, and R. Yao. The Termination of Commercial Mortgage Contracts through Prepayment and Default: A Proportional Hazard Approach with Competing Risks. *Real Estate Economics*, 2002, 30:4, 595–633.

Clapp, J.M., Y. Deng, and X. An. Unobserved Heterogeneity in Models of Competing Mortgage Termination Risks. *Real Estate Economics*, 2006, 34:2, 243–73.

Consumer Financial Protection Bureau. What is a Qualified Mortgage? 2013.

——. General Comparison of Ability to Repay Requirements with Qualified Mortgages. 2014.

Demyanyk, Y. and O. Van Hemert. Understanding the Subprime Mortgage Crisis. *Review of Financial Studies*, 2011, 24:6, 1848–80.

Deng, Y. Mortgage Termination: An Empirical Hazard Model with a Stochastic Term Structure. *Journal of Real Estate Finance and Economics*, 1997, 14:3, 309–31.

Elul, R., N.S. Souleles, S. Chomsisengphet, D. Glennon, and R. Hunt. What ''Triggers'' Mortgage Default. *American Economic Review*, 2010, 100:2, 490–94.

Fair Isaac Corporation. How My FICO Scores are Calculated. myFico.com. Available at: http://www.myfico.com/crediteducation/whatsinyourscore.aspx, 2015.

Fannie Mae. Fannie Mae Annual Report 2007. Available at: http://www.fanniemae.com/resources/file/ir/pdf/proxy-statements/2007_annual_report.pdf, 2008.

——. 97% LTV Options for Purchases and Limited Cash-Out Refinance of Fannie Mae Loans. Available at: https://www.fanniemae.com/content/faq/97-ltv-options-faqs.pdf, 2015.

Fayyad, U., G. Piatetsky-Shapiro, and P. Smyth. From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, 1996, 17:3, 37.

Federal Reserve. Mortgage Debt Outstanding. Available at: http://www.federalreserve.gov/econresdata/releases/mortoutstand/current.htm, 2016.

Feldman, D. and S. Gross. Mortgage Default: Classification Trees Analysis. *Journal of Real Estate Finance and Economics*, 2005, 30:4, 369–96.

Financial Crisis Inquiry Commission. The Financial Crisis Inquiry Report. U.S. Government Printing Office, 2011.

Foote, C.L., K. Gerardi, and P.S.Willen. Negative Equity and Foreclosure: Theory and Evidence. *Journal of Urban Economics*, 2008, 64:2, 234–45.

Freddie Mac. Freddie Mac Annual Report 2007. Available at: http://www.freddiemac.com/investors/ar/pdf/2007annualrpt.pdf, 2008.

Gerardi, K., L. Goette, and S. Meier. Financial Literacy and Subprime Mortgage Delinquency: Evidence from a Survey Matched to Administrative Data. Federal Reserve Bank of Atlanta Working Papers, September, 2010.

Ghatasheh, N. Business Analytics using Random Forest Trees for Credit Risk Prediction: A Comparison Study. *International Journal of Advanced Science and Technology*, 2014, 72, 19–30.

Hosmer, Jr., D.W. and S. Lemeshow. *Applied Logistic Regression.* John Wiley & Sons, 2004.

Jiang, W., A.A. Nelson, and E. Vytlacil. Liar's Loan? Effects of Origination Channel and Information Falsification on Mortgage Delinquency. *Review of Economics and Statistics*, 2014, 96:1, 1–18.

Kan, R. and C. Robotti. The 2008 Federal Intervention to Stabilize Fannie Mae and Freddie Mac, 2007.

Kohavi, R. A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection. *IJCAI*, 1995, 1137–45.

Lacour-Little, M., Y.W. Park, and R.K. Green. Parameter Stability and the Valuation of Mortgages and Mortgage-Backed Securities. *Real Estate Economics*, 2012, 40:1, 23–63.

LaCour-Little, M. and J. Yang. Taking the Lie Out Of Liar Loans: The Effect of Reduced Documentation on the Performance and Pricing of Alt-A and Subprime Mortgages. *Journal of Real Estate Research*, 2013, 35:4, 507–53.

Mayer, C., K. Pence, and S.M. Sherlund. The Rise in Mortgage Defaults. *The Journal of Economic Perspectives*, 2009, 23:1, 27–50.

Park, Y.W. and D.W. Bang. Loss Given Default of Residential Mortgages in a Low LTV Regime: Role of Foreclosure Auction Process and Housing Market Cycles. *Journal of Banking and Finance*, 2014, 39:1, 192–210.

Peni, E., Smith, S. and S. Vähämaa. Bank Corporate Governance and Real Estate Lending During the Financial Crisis. *Journal of Real Estate Research*, 2013, 35:3, 313–43.

Peterson, C.L. Fannie Mae, Freddie Mac, and the Home Mortgage Foreclosure Crisis. *Loyola University New Orleans Journal of Public Interest Law*, 2008, 149–70.

Qi, Z., Y. Tian, and Y. Shi. Laplacian Twin Support Vector Machine for Semi-supervised Classification. *Neural Networks*, 2012, 35, 46–53.

Quercia, R.G. and M. Stegman. Residential Mortgage Default: A Review of the Literature. *Journal of Housing Research*, 1992, 3:2, 341–80.

R Project. The R Project for Statistical Computing. Available at: http://www.r-project.org/.

Reuters, 2008. Fannie Mae Tightens Loan Standard to Protect Itself. *The New York Times*. Available at: http://www.nytimes.com/2008/04/02/business/rtlend-web.html, 2015.

Safavian, S.R. and D. Landgrebe. A Survey of Decision Tree Classifier Methodology. *IEEE Transactions on Systems, Man, and Cybernetics*, 1991, 21:3, 660–74.

Schmeiser, M.D. and M.B. Gross. The Determinants of Subprime Mortgage Performance Following a Loan Modification. *Journal of Real Estate Finance and Economics*, 2016, 52: 1, 1–27.

Shenn, J. Fannie Mae Tightens Mortgage Standards for Some Home Buyers. *BloombergBusiness*, 2012.

Sherlund, S.M. *Mortgage Defaults*. 2010.

Shiller, R.J. *Irrational Exuberance*. Princeton University Press, 2015.

Smith, B.C. Stability in Consumer Credit Scores: Level and Direction of FICO Score Drift as a Precursor to Mortgage Default and Prepayment. *Journal of Housing Economics*, 2011, 20:4, 285–98.

Spahr, R. and M. Sunderman. The U.S. Housing Finance Debacle, Measures to Assure its Non-recurrence, and Reform of the Housing GSEs. *Journal of Real Estate Research*, 2014, 36:1, 59–86.

Sun, Z. *Classification System for Mortgage Arrear Management*. University of Groningen, 2013.

Wallison, P.J. and C.W. Calomiris. The Last Trillion-Dollar Commitment: The Destruction of Fannie Mae and Freddie Mac. *Journal of Structured Finance*, 2009, 15:1, 71–80.

Yegnanarayana, B. *Artificial Neural Networks*. PHI Learning Pvt. Ltd., 2009.

Young, J.T. The Worst Four Years of GDP Growth In History: Yes, We Should Be Worried. *Forbes*, 2013.

Zivot, E. and D.W.K. Andrews. Further Evidence on the Great Crash, the Oil-price Shock, and the Unit-root Hypothesis. *Journal of Business & Economic Statistics*, 2002, 20:1, 25–44.

Zurada, J., N. Kunene, and J. Guan. The Classification Performance of Multiple Methods and Datasets: Cases from the Loan Credit Scoring Domain. *Journal of International Technology and Information Management*, 2014, 23:1, 57–82.

Stanislav Mamonov, Montclair State University, Montclair, NJ 07043 or stanislav.mamonov@montclair.edu.

Raquel Benbunan-Fich, Baruch College, CUNY, New York, NY 10010 or rbfich@baruch.cuny.edu.