Bayesian Forecasting of Prepayment Rates for Individual Pools of Mortgages

Ivilina Popova
Assistant Professor, Seattle University

Elmira Popova
Associate Professor, The University of Texas at Austin

Edward George
Professor, University of Pennsylvania

Ivilina Popova
Department of Finance, Albers School of Business and Economics, Seattle University, Seattle, WA 98122, USA, popovai@seattleu.edu

Elmira Popova
Graduate Program in Operations Research and Industrial Engineering, The University of Texas at Austin, Austin, TX 78712, USA, elmira@mail.utexas.edu

Edward George
Department of Statistics, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104, USA, edgeorge@wharton.upenn.edu

JEL Classification:G, G21, C5, C11, C15
Keywords: mortgages, prepayment rates, Bayesian forecasting

This research has been partially supported by grant #003658-0763 from the State of Texas Advanced Research Program and NSF grant #DMI-0457558

Bayesian Forecasting of Prepayment Rates for Individual Pools of Mortgages

Abstract

This paper proposes a novel approach for modeling prepayment rates of individual pools of mortgages. The model incorporates the empirical evidence that prepayment is past dependent via Bayesian methodology. There are many factors that influence the prepayment behavior and for many of them there is no available (or impossible to gather) information. We implement this issue by creating a Bayesian mixture model and construct a Markov Chain Monte Carlo algorithm to estimate the parameters. We assess the model on a data set from the Bloomberg Database. Our results show that the burnout effect is a significant variable for explaining normal prepayment activities. This result does not hold when prepayment is trigerred by non-pool dependent events.

1 Introduction

Purchasing a house usually involves obtaining a loan (mortgage) originated by a financial institution. Any standard mortgage monthly payment consists of scheduled interest and principal. The borrower is also allowed to include additional payment toward the principal or early payoff the whole mortgage. Refinancing of the mortgage is an example involving such a prepayment. The issuer of the mortgage usually sells the mortgages to another financial institution that pools them together and issues new securities, commonly known as mortgage backed securities. The buyers of such structured products would like to know in advance the size of the incoming prepayments (if any). To obtain the correct price one needs to know a forecast of the prepayment.

One possible model of the prepayment is if we pretend that the borrower holds a call option on the loan with exercise price equal to the outstanding balance (i.e. it is a time varying strike). Under optimal exercising conditions a mortgage can be priced as a callable bond. One would expect that the holder will prepay when the refinancing rate is below the mortgage rate. However, the empirical evidence, which shows very different behavior [7], does not support such a model. As has been documented, homeowners often prepay when it is not optimal to do so and vice versa.

The recent literature attempts to model this partially "irrational" behavior. The prepayment activities can be classified as either interest rate related or non-interest rate related. The interest rate related activities (optimal prepayment) occur when the homeowners act in order to minimize the market value of the loan. The non-interest rate activities (suboptimal prepayments) occur for personal reasons of the borrowers, such as divorce, job change, etc.

Two approaches have been considered for modeling prepayment. Dunn and McConnell [7] pioneered a model based on standard contingent claim pricing theory. In their model the prices of the mortgage backed securities (MBS) and prepayment behavior are determined together. They assume Cox-Ingersoll-Ross interest rate model [3], introduce a suboptimal prepayment as a Poisson event, and using non-arbitrage argument derive a partial differential equation for the price of the security that can be solved numerically. Suboptimal prepayment behavior was first documented in this study.

A second approach is empirical - based on historical information. Schwartz and Torous [12] model the prepayment rate as a function of different explanatory variables, like seasonality, burnout, difference between the contracting and re financing rates, and speed of prepayment. Again using a standard arbitrage argument they derive a partial differential equation that the MBS need to satisfy.

Richard and Roll [11] describe the prepayment model used by Goldman Sachs. They consider four important effects: refinancing incentive, age of the mortgage, seasonality and burnout. The conditional prepayment rate is determined as a product of functions of the above factors. A non-linear least squares optimization procedure gives the estimated values of the parameters. Stanton [14] presents a model that extends the option-theoretic approach. He models the transaction costs faced by mortgage holders and assumes that prepayment decisions occur at discrete time. This produces prepayment behavior that is consistent with the burnout effect.

All of the existing models estimate the prepayment function by using the information from all pools in the sample. In other words, they assume that all the pools manifest similar prepayment behavior. Recently, Stanton [15], investigates the problem of predicting the prepayment for individual pools of mortgages. He reports that in 1,000 GNMA mortgage pools over a six and one-half year period, the unobservable heterogeneity is statistically significant. This could lead to very different prices of MBS that are backed by different pools. One of the latest models used by Wall Street firms (BlackRock) predicts individual pool prepayment rates based on detailed information about individual loans in the pools.

None of the previous models has dealt with pool level predictions, and this is our main research objective. Section §2 describes the nature of the raw data and how the final data set is constructed. Section §3 describes the statistical prepayment model, Section §4 presents the empirical results and Section §5 concludes.

2 Prepayment Data Description

We consider historical data for 74 pools of mortgages, split between issues of Freddie Mac[1] and Ginnie Mae[2]. To ensure the homogeneity of the data, we exclusively focused on 30 year fixed rate single family 8% coupon mortgages. For these pools, Bloomberg provided general and monthly pool information consisting of: issue date, maturity date, original amount, historical monthly prepayment (as percentage PSA). PSA stands for the Public Securities Association convention which assumes that 0.2% of the principal is paid in the first month, increase by 0.2% for the following 29 months, and flattens at 6% until maturity. We chose pools that have at least $1 million original amount. The age of the mortgages varies from 5 to 25 years, and consequently, the number of data points for each pool from 60 to 300. Tables I and II show the CUSIPs (a security identifier as defined by the Committee on Uniform Securities Identification Procedure) of the pools, their initiation date, initial dollar amount and current age (in months).

INSERT TABLE I AND II HERE

The average age for Freddie Mac pools is 122 months and the average initial dollar amount is approximately $9 million. The average age for Ginnie Mae pools is 300 months and the average initial dollar amount is approximately $3 million. In addition to the information provided by Bloomberg, we gathered historical long (30 year Treasury Bond) and short term (3 month Treasury Bill) interest rates from [8].

The construction of the data set consists of several steps. First, we computed the conditional prepayment rate (CPR) using the formula: $CPR = PSA(standard).PSA(historical)/100$, where $PSA(historical)$ comes from the data set. Next, we computed the single monthly mortality rate as $SMM = 1 - (1 - CPR)^{\frac{1}{12}}$. From here we can compute the monthly payment,

[1]Stockholder-owned corporation chartered by U. S. Congress to increase the supply of money that mortgage lenders can make available to homebuyers and multifamily investors.

[2]The U. S. Government National Mortgage Association

MP, and the interest payment, IP: For $t = 0, \ldots, 360$

$$MP = \frac{FaceValue\left[\frac{0.08}{12}\right]\left[1+\frac{0.08}{12}\right]^t}{\left[1+\frac{0.08}{12}\right]^t - 1}$$

and

$$IP = FaceValue\left[\frac{0.08}{12}\right]$$

Then the scheduled principal will equal to, $SP = MP - IP$, the nonscheduled prepayment, $NPP = SMM(FaceValue - SP)$, and the actual payment is $AP = SP + NNP$.

In our analysis we model the actual payment (in dollar amount) from a mortgage pool $i = 1, \ldots, 74$ at the end of each month t and denote it by AP_t^i.

To assess the general structure of the data and to identify the presence of multimodality we used kernel density estimation, see [13]. We entered as an input the natural logarithm of the actual monthly payments for all 74 pools of mortgages. The produced density estimators showed that the majority of the mortgage pools have at least 2 modes.

Insert Figure 1 here

Figure 1 shows 9 such densities[3]. Since we model the total payment from a pool i, the kernel density estimators are based on the N_i monthly total payments from that pool. The observed multimodality is intuitively reasonable - it seems likely that there are borrowers who prepay small amounts each month, as well as borrowers who prepay the whole mortgage (refinancing, selling the house, etc.) There are probably certain events that trigger the "small" or the "large" prepayment behavior. If we knew these events and could gather data associated with them, then we might be able to predict the next month prepayment. However, up to now, there is no research regarding such events, their existence and data availability.

[3] All 74 densities are available upon request.

3 Bayesian Prepayment Model

The plots in Figure 1 suggest that the distribution of AP_t^i is not unimodal and may be better described with a bimodal model such as a mixture of normals. In particular, we model the AP_t^i values as the realizations of N_i independent random variables (N_i monthly total payments in the pool i), $x_t^i, t = 1, \ldots, N_i$ from a 2-component mixture

$$f\left(x_t^i\right) = p^i f_1(x_t^i) + (1 - p^i) f_2(x_t^i), t = 1, \ldots, N_i$$

where $f_1(.)$ is the probability density function of a Normal distribution with parameters μ_1^i and σ_1^i, and $f_2(.)$ is the probability density function of a Normal distribution with parameters μ_2^i and σ_2^i. We denote the unknown parameters of the mixture by $\theta^i = \left(\mu_1^i, \mu_2^i, (\sigma_1^i)^2, (\sigma_2^i)^2, p^i\right), i = 1, 2, \ldots, 74$.

To fit and draw inference from such a model, we take a Bayesian approach which describes the uncertainty of the parameters θ^i with prior probability distributions.

Dempster, Laird and Rubin [4] show that any mixture model can be expressed in terms of *missing* or *incomplete* data as follows. Define for pool i and $1 \leq t \leq N_i$, z_t^i as a 2 - dimensional random vector indicating to which component x_t^i belongs, i.e. $z_{tj}^i \in \{0, 1\}$, $t = 1, \ldots, N_i$, $j = 1, 2$. z_t^i's are independent for $t = 1, \ldots, N_i$ and the density of the completed data $\left(x_t^i, z_t^i\right)$ for pool i is

$$\prod_{j=1}^{2} (p_j^i)^{z_{tj}^i} f_j^{z_{tj}^i}\left(x_t^i\right), t = 1, \ldots, N_i$$

Using this representation, the parameters θ^i, of the mixture model can be estimated by a Markov Chain Monte Carlo (MCMC) algorithm for each i, $i = 1, \ldots, 74$. The general description of such an algorithm can be found in [9].

The preliminary analysis of the data suggests that there are two main groups - "small" and "large" prepayers. We would like to model this structure by also including explanatory variables (covariates) found to be significant for predicting prepayment behavior in the current literature. Schwartz and Torous [12] model time until prepayment of a mortgage using a proportional hazard model. In their model the effect of the covariates is on the hazard function which they refer as prepayment function. (The hazard function at time t is the probability that a prepayment will occur in time $t + \triangle t$ given no prepayment occurred by time t). The variables described bellow are first defined in their paper.

For each pool i at time t define the following variables:

- $X_1^i(t)$ = the difference between the mortgage rate and the short term interest rate
- $X_2^i(t) = \left(X_1^i\right)^3$
- $X_3^i(t)$ captures the burnout effect - it is the logarithm of the ratio of the dollar amount of the pool i outstanding at time t, to the pool's principal which would prevail at t in the absence of prepayments
- $X_4^i(t)$ models the seasonality effect. It equals 1 for the months of May, June, July, and August, and 0 for September, October, November, December, January, February, March, and April.
- $X_5^i(t)$ is the spread (difference) between the long and short term interest rates.

As a forecasting model we define a Bayesian mixture of regressions (see [10] for an alternative presentation), where $Y^i(t) = \ln(AP_t^i)$ is the logarithm of total money paid at the end of the month t for pool i, and $X_1^i(t), X_2^i(t), X_3^i(t), X_4^i(t), X_5^i(t)$ are the covariates defined above.

The model is:

$$Y^i(t) \sim p^i N\left(\sum_{k=0}^{5} U_k^i X_k^i(t), W_1^i\right) + (1 - p^i) N\left(\sum_{k=0}^{5} V_k^i X_k^i(t), W_2^i\right)$$

where $X_0^i(t) = 1$ for all $t = 1, \ldots N_i, i = 1, \ldots, 74$. Note that we use W_1^i and W_2^i to represent the precision parameters, rather than the variances, of the two normal distributions that constitute the mixture[4].

[4] $Precision = 1/Variance$

Letting $M_1^i = (U_0^i, \ldots, U_5^i)$ and $M_2^i = (V_0^i, \ldots, V_5^i)$ to denote covariate coefficients corresponding to each of the mixture components, we consider the following default priors for the unknown parameters. For a fixed i we assumed improper joint prior distribution for (M_j^i, W_j^i), $\xi(m_j^i, w_j^i) = 1/w_j^i, w_j^i > 0$ for $j = 1, 2$. We should point out that the prior distributions for (M_1^i, W_1^i) and (M_2^i, W_2^i) are independent. Assume Beta(α^i, β^i) as a prior distribution for p^i.

Define X^i to be the matrix

$$X^i = \left\{ \begin{array}{llllll} 1 & X_1^i(1) & X_2^i(1) & X_3^i(1) & X_4^i(1) & X_5^i(1) \\ 1 & X_1^i(2) & X_2^i(2) & X_3^i(3) & X_4^i(2) & X_5^i(2) \\ \cdots & & & & & \\ 1 & X_1^i(N_i) & X_2^i(N_i) & X_3^i(N_i) & X_4^i(N_i) & X_5^i(N_i) \end{array} \right\},$$

$$Y^i = \Big(Y^i(1), Y^i(2), \ldots, Y^i(N_i)\Big)'$$

and

$$\hat{m}^i = \left[(X^i)' X^i\right]^{-1} (X^i)' Y^i \qquad (1)$$

$$(\hat{s}^i)^2 = \frac{1}{N_i - 2} \left(Y^i - X^i \hat{m}^i\right)' \left(Y^i - X^i \hat{m}^i\right) \qquad (2)$$

$$n_1^i = \text{Number of data points in cluster 1 from pool } i$$

$$n_2^i = \text{Number of data points in cluster 2 from pool } i$$

$$n_1^i + n_2^i = N_i$$

Under our priors, the full conditional posterior distributions of the mixture parameters are (see [6]):

$$M_j^i \sim \text{Multivariate } t\left(n_j^i - 2, \hat{m}^i, \frac{1}{(\hat{s}^i)^2} (X^i)' X^i\right) \qquad (3)$$

$$W_j^i \sim \text{Gamma}\left(\frac{n_j^i - 2}{2}, \frac{n_j^i - 2}{2} (\hat{s^i})^2\right)$$

$$p^i \sim \text{Beta}\left(n_1^i + \alpha^i, n_2^i + \beta^i\right)$$

$$j = 1, 2$$

The Gibbs sampling algorithm (ran separately for each i) consists of the following steps:

1. Start with initial values of the parameters $(\theta^i)^0 = \left[(M_1^i)^0, (M_2^i)^0, (W_1^i)^0, (W_2^i)^0, (p^i)^0\right]$

2. Allocate each $Y^i(t)$ to the first or second mixture based on odds ratio

$$\frac{P\left[z_t^i = 0|\theta^i, Y^i(t)\right]}{P\left[z_t^i = 1|\theta^i, Y^i(t)\right]}$$

   This part of the algorithm will generate $z_t^i = 0$ if the odds ratio is less than 1 and $z_t^i = 1$ if it is greater than 1.

3. As a result, each of the monthly payments gets assigned either to cluster 1 or to cluster 2. Given that we are in cluster j, we simulate θ^i from the marginal distributions (3). Note that conditional on z_t^i's :M_1^i and M_2^i are independent, and W_1^i and W_2^i are independent.

4. Repeat the above steps L times (number of simulation runs).

4 Empirical Results

We implemented the above algorithm and used the uniform random number generator written by P. L'Ecuyer (see [17]) and non uniform random number algorithms from Devroye, [5]. The length of our simulation runs is $L = 5000$, the algorithm is implemented in C++ and run on 2.37GHz Pentium. As with any MCMC algorithm, there is a transient period at the beginning of the simulation. Figure 2 shows the actual simulated points for U_0, V_0, W_1 and W_2 for one of the pools of mortgages. Based on the observed patterns from the time trace plot of the simulated values we dropped the first 1000 simulated values and used only the last 4000 simulated values for estimation and inference. We observed simular patterns with the remaining 73 pools of mortgages.

Insert Figure 2 here

Before presenting our model estimates, we emphasize that we have treated all the parameters as random variables and the estimates that we use are the expected values of the corresponding posterior distributions. In the Bayesian approach we can construct the true probability intervals around the parameter expected values. In the classical approach we usually have the estimate and we compute its statistical significance or construct a confidence interval. Here we work with a whole distribution of values instead of one parameter estimate. We computed 95% probability intervals for all parameter values. Table 3 shows results for all Freddie Mac pools and Table 4 shows results for all Ginnie Mae pools. We reported only the parameters with probability intervals not covering zero for at least 50% of the pools[5]. For Freddie Mac the relevant parameters are U_0, V_0 and V_3, where U_0 and V_0 are the intercepts and V_3 is the variable measuring the burnout effect. For Ginnie Mae the parameters are U_0, U_3, V_0 and V_3, where U_0 and V_0 are the intercepts and U_3 and V_3 measure the burnout effect.

Insert Tables III and IV here

Note that most of the coefficients for V_3 and U_3 are negative and the corresponding probability intervals do not include zero. This result is very intuitive since as the prepaid amount increases the likelihood that there will be a prepayment decreases. There are only two Freddie Mac pools that have positive burnout coefficients. After examining their prepayment history we observed that the prepayment activity increased as the burnout effect increased, so the results are consistent with the historical prepayment behavior.

Insert Table V here

Table V summarize the results from Tables III and IV. We report the average parameter values across all pools. For Ginnie Mae pools the average burnout coefficient for the first cluster is -0.62 and for the second cluster is -1.23 with an average probability of 42% of choosing cluster one and 58% of choosing cluster two. The corresponding results for the Freddie Mac pools are -0.94 for the burnout coefficient of the second cluster with an average probability of

[5] If the probability interval includes zero, we cannot reject the hypothesis that the true parameter value is different from zero, therefore it may not be significant.

48% of choosing cluster one and 52% of choosing cluster 2. The common theme here is that the coefficient is "more negative" for the distribution that occurs more often, which represents the "normal prepayment behavior". For normal prepayment activities it is to be expected that the prepayment will decrease as the principal amount left in the pool decreases. In contrast, the second distribution represents prepayment that are out-of-the-norm, they can be triggered by non-pool related variables such as selling the house, moving, bankruptcy, etc.

Given that we were able to construct and compute the expected posterior values for each of the model's parameters, we examine the forecast quality for each of the pools and for all pools together. Figure 3 compares the actual prepayment with the predicted prepayment for one pool of mortgages.

INSERT FIGURE 3 HERE

The lower solid curve is $\sum_{k=0}^{5} \hat{U}_k X_k(t)$, where $\hat{U}_k$ are the expected values of the parameters posterior distributions. The upper solid curve is $\sum_{k=0}^{5} \hat{V}_k X_k(t)$, where $\hat{V}_k$ are the expected values of the parameters posterior distributions. The model will forecast an "average" prepayment with probability $\hat{p}$ and a "high" prepayment with probability $1 - \hat{p}$. A point forecast could be the weighted average of these two points (the middle line). We use that forecast from the mixture to compute the associated error (residual).

For every pool of mortgages we compute the mean and the standard deviation of the associated residuals. We use the computed 74 means and standard deviations to construct the histograms of the calculated means and standard deviations.

INSERT FIGURE 4 AND 5 HERE

Figure 4 shows the histogram of the means of the residuals (measured as the actual minus the predicted using the means of the posterior distributions as estimates for the model parameters). Figure 5 shows the histogram of the standard deviations of the residuals. The mean of the means of the residuals is -0.07228 and the mean of the standard deviation of the residuals is

0.62463. Note that both histograms have large right tails due to results of one Freddie Mac pool that has one extremely large prepayment point[6]. Next we performed a t-test on the computed 74 means of the residuals. The p-value of the one sample t-test is 0.17 and the test fail to reject the hypothesis the true mean is equal to zero. This result indicates that the model has a good predictive power.

5 Conclusion

In this paper, we have proposed using two component normal regression mixture models to describe the apparent bimodal distribution of prepayments within mortgage pools over time. Distinct linear regression functions of observed covariates are used to model the means of the two components. We consider a model where the mixture probabilities are constrained to be fixed over time. To fit this multiparameter nonlinear model, we take a Bayesian statistical approach where the uncertainty about all the unknown parameters is described by prior distributions. For this setup, an MCMC algorithm is constructed and used to carry out all the computations. Empirical results show that the fixed mixture weight model appears to fit the observed data reasonably well.

We feel that our model is a good start towards the modeling of the individual pool prepayment rate process. Naturally, the extent to which this model can be effective depends on the quality of the available covariates. More and better covariate information will likely lead to better fits and forecasts. Another future direction for potential improvements will be to consider elaborations to larger models that can exploit information across similar mortgage pools. One such elaboration would be a hierarchical Bayes model that treats the parameters of each pool as a sample from a superpopulation model. Such an elaboration would be particularly natural given the Bayesian treatment we have here considered.

[6]Unfortunately we cannot check if this large prepayment is due to a data error or it is a real value. We assumed that it is a real observation.

References

[1] Albert J. H. and S. Chib, Bayesian Analysis of Binary and Polychotomous Response Data, *Journal of the American Statistical Association*, 88 (422), 669-679, 1993.

[2] Bernardo J. M. and A. F. M. Smith, *Bayesian Theory*, John Wiley and Sons, 1993.

[3] Cox J. C., Ingersoll J. E. Jr. and S. A. Ross. A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2):385-407, 1985.

[4] Dempster, A. P., Laird, N. M. and Rubin, D. B., Maximum Likelihood from Incomplete Data via the EM algorithm (with discussion), *J. R. Statist. Soc.*, B, 39, 1-38, 1977.

[5] Devroye Luc, *Non-Uniform Random Variate Generation*, `http://cgm.cs.mcgill.ca/~luc/rnbookindex.html`

[6] Diebolt J. and C. P. Robert, Estimation of Finite Mixture Distributions through Bayesian Sampling, *J. R. Statist. Soc.*, B, 56, No. 2, 363-375, 1994.

[7] Dunn K. and J. McConnell, Valuation of GNMA Mortgage-Backed Securities, *The Journal of Finance*, vol. XXXVI, No. 3, 599-616, 1981.

[8] `http://www.stld.frb.org/fred/`

[9] Gilks W. R., Richardson S. and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*, Chapman and Hall, 1996.

[10] Hurn M., Justel A. and C. P. Robert, Estimating Mixtures of Regressions, *J. Comput. Graph. Stat.*, 12(1), 1-25, 2003.

[11] Richard R. and R. Roll, Modeling Prepayments on Fixed-Rate Mortgage-Backed Securities, *Journal of Portfolio Management*, 15 (3), 73-82, 1989.

[12] Schwartz E. and W. Torous, Prepayment and the Valuation of Mortgage-Backed Securities, *Journal of Finance*, 44(2), 375-392, 1989.

[13] Silverman B. W., Using Kernel Density Estimates to Investigate Multimodality, *J. R. Statist. Soc.*, 43, No. 1, pp. 97-99, 1981.

[14] Stanton R., Rational Prepayment and the Valuation of Mortgage-Backed Securities, *Review of Financial Studies*, Vol. 8, No. 3, 677-708, 1996a.

[15] Stanton R., Unobservable Heterogeneity and Rational Learning: Pool-Specific versus Generic Mortgage-Backed Security Prices, *Journal of Real Estate and Economics*, 12: 243-263, 1996b.

[16] Zellner A., Optimal information processing and Bayes' theorem, *American Statistician*, 42, 278-284, 1988.

[17] `http://www.iro.umontreal.ca/~lecuyer`

Table I
Ginnie Mae Pool Information

Tables I shows the CUSIPs (a security identifier as defined by the Committee on Uniform Securities Identification Procedure) of the Ginie Mae pools, their initiation date, initial dollar amount and current age (in months).

CUSIP	Origination Date	Initial Amount		Age in months
362027BX1	1/1/1973	$	5,388,014.99	308
362027JP0	1/1/1973	$	2,303,787.79	308
362027J47	1/1/1973	$	2,156,245.31	308
362027KB9	1/1/1973	$	1,994,981.34	308
362026VD5	2/1/1973	$	4,057,248.04	304
362027RC0	4/1/1973	$	4,199,561.59	305
362027R55	4/1/1973	$	1,709,405.67	305
362027SS4	4/1/1973	$	3,203,047.00	305
362027UJ1	6/1/1973	$	3,040,361.00	304
362027UK8	6/1/1973	$	3,812,288.98	304
362027V76	7/1/1973	$	2,818,570.48	303
362027FQ2	11/1/1973	$	2,249,925.76	298
362027HZ0	11/1/1973	$	2,029,476.62	299
362027KW3	11/1/1973	$	2,015,190.05	299
362027N26	11/1/1973	$	3,642,915.05	299
3620274Q4	11/1/1973	$	2,001,116.95	297
362028A77	11/1/1973	$	2,020,035.56	298
362028BT8	11/1/1973	$	2,007,416.51	299
362028FM9	11/1/1973	$	5,000,616.25	299
362027HS6	12/1/1973	$	2,020,564.41	298
362027QY3	12/1/1973	$	2,010,207.98	296
362027T95	12/1/1973	$	7,912,679.24	298
3620274L5	12/1/1973	$	2,015,578.53	298
3620274N1	12/1/1973	$	2,194,180.55	298
3620276T6	12/1/1973	$	3,011,643.62	298
362028AZ5	12/1/1973	$	5,000,764.88	298
362028BX9	12/1/1973	$	2,008,693.49	298
362028DY5	12/1/1973	$	5,009,615.27	298
362028EQ1	12/1/1973	$	10,234,823.28	298
362028FB3	12/1/1973	$	4,021,388.17	297
362028F31	12/1/1973	$	2,000,630.91	296
362028HB1	12/1/1973	$	2,014,895.05	298
362028HU9	12/1/1973	$	2,000,041.48	298
362028JF0	12/1/1973	$	2,499,138.92	298
362028JK9	12/1/1973	$	4,017,510.77	297

Table II
Freddie Mac Pool Information

Table II shows the CUSIPs (a security identifier as defined by the Committee on Uniform Securities Identification Procedure) of the Freddie Mac pools, their initiation date, initial dollar amount and current age (in months).

CUSIP	Origination Date	Initial Amount	Age in months
31340B5T2	12/1/1988	$4,022,391.00	118
31340B5Y1	12/1/1988	$3,425,491.00	118
31340CBW6	11/1/1992	$33,317,099.00	71
31340CB55	5/1/1993	$9,342,593.00	65
31340CEJ2	9/1/1989	$93,073,150.00	109
31340CEL7	9/1/1989	$1,199,680.00	109
31340CER4	9/1/1989	$4,111,316.00	109
31340CE60	10/1/1989	$1,312,970.00	108
31340CFR3	12/1/1989	$3,012,953.00	106
31340CBK2	11/1/1991	$33,884,130.00	83
31340CBL0	12/1/1991	$12,484,193.00	82
31340CBP1	1/1/1992	$14,765,342.00	81
31340CBR7	2/1/1992	$5,805,280.00	80
31340CBS5	2/1/1992	$9,970,966.00	80
31340CBA4	7/1/1989	$20,690,305.00	111
31340CBG1	9/1/1991	$20,843,194.00	85
31340CBJ5	10/1/1991	$28,367,731.00	84
31340CAH0	3/1/1989	$1,068,345.00	115
31340CAJ6	3/1/1989	$4,053,510.00	115
31340CAN7	3/1/1989	$3,157,634.00	115
31340B6B0	1/1/1989	$3,029,407.00	117
31340B6J3	1/1/1989	$1,000,811.00	117
31340B6V6	2/1/1989	$1,198,354.00	116
31340B7L7	2/1/1989	$1,472,666.00	116
31340CAA5	2/1/1989	$1,007,985.00	116
31340CAC1	2/1/1989	$1,234,631.00	116
31340CAF4	3/1/1989	$1,711,867.00	115
31340B5P0	12/1/1988	$1,209,824.00	118
31340B5H8	12/1/1988	$2,191,105.00	118
31340AEU1	7/1/1984	$3,924,740.00	170
31340AEL1	6/1/1984	$3,328,475.00	171
31340AD40	4/1/1984	$1,687,252.00	173
31340ADW8	3/1/1984	$2,856,860.00	174
31340ADH1	1/1/1984	$9,493,396.00	174
31340ADG3	1/1/1984	$8,615,957.00	176
31340ADF5	1/1/1984	$1,081,768.00	176
31340ADC2	12/1/1983	$1,031,959.00	177
31340AC82	12/1/1983	$1,443,159.00	177
31340ABB6	9/1/1983	$1,354,125.00	180

Table III
Estimated values for the Parameters of the Bayesian Mixture of Regressions for Freddie Mac Pools

Table III shows the estimated values for model parameters, their standard deviation and 95% Probability intervals. Results are reported for all Freddie Mac mortgage pools. For example, for pool with CUSIP = 31340ABB6, the estimated value of the intercept is 9.66 with a standard deviation of 0.05. LPL stands for Left Probability Limit and RPL stands for Right Probability Limit. U_0 and V_0 are the intercepts and V_3 is the burnout effect coefficient of the regression model that gets chosen with probability 1-p.

	PARAMETER U_0				PARAMETER V_0				PARAMETER V_3			
Pool CUSIP	Average	SD	LPL	RPL	Average	SD	LPL	RPL	Average	SD	LPL	RPL
31340abb6	9.66	0.05	9.57	9.75	10.87	0.08	10.72	11.02	-1.26	0.03	-1.32	-1.20
31340ac82	9.29	0.11	9.07	9.50	11.40	0.45	10.52	12.28	-0.81	0.24	-1.28	-0.33
31340ad40	9.95	0.19	9.57	10.33	9.88	0.21	9.48	10.28	-1.58	0.16	-1.89	-1.28
31340adc2	9.13	0.96	7.26	11.01	9.24	0.39	8.48	10.00	-1.21	0.53	-2.25	-0.16
31340adf5	9.46	0.29	8.90	10.03	9.92	0.43	9.08	10.77	-1.81	0.36	-2.50	-1.11
31340adg3	11.19	0.05	11.09	11.29	11.78	0.14	11.50	12.06	-0.81	0.11	-1.02	-0.59
31340adh1	11.59	0.14	11.31	11.86	11.15	0.19	10.77	11.53	-1.03	0.12	-1.27	-0.79
31340adw8	9.90	0.14	9.63	10.17	11.40	0.33	10.75	12.05	-1.26	0.07	-1.40	-1.11
31340ael1	10.13	0.40	9.34	10.92	11.27	0.35	10.58	11.96	-1.22	0.09	-1.39	-1.05
31340aeu1	10.33	0.04	10.25	10.42	11.01	0.03	10.95	11.08	-1.25	0.05	-1.36	-1.14
31340b5h8	9.77	0.05	9.67	9.87	11.09	0.31	10.47	11.70	-0.56	0.14	-0.84	-0.27
31340b5p0	9.62	0.13	9.37	9.88	9.26	0.12	9.02	9.51	-1.77	0.70	-3.14	-0.39
31340b5t2	10.98	0.42	10.15	11.81	10.59	0.49	9.62	11.55	-0.84	0.47	-1.75	0.07
31340b5v7	10.10	0.08	9.95	10.25	1.50	3.34	-5.05	8.05	3.26	0.35	2.57	3.94
31340b5y1	10.17	0.10	9.98	10.35	10.40	0.26	9.89	10.91	-1.58	0.34	-2.26	-0.91
31340b6b0	10.06	0.09	9.90	10.23	10.21	0.34	9.55	10.88	-1.14	0.10	-1.34	-0.94
31340b6j3	8.96	0.10	8.77	9.15	8.00	0.55	6.92	9.09	-1.15	0.40	-1.94	-0.36
31340b6v6	9.27	0.09	9.10	9.44	9.72	0.31	9.11	10.32	-1.28	0.23	-1.73	-0.83
31340b7l7	9.36	0.08	9.20	9.51	10.86	0.45	9.98	11.73	-1.03	0.28	-1.57	-0.49
31340caa5	8.82	0.05	8.72	8.91	9.57	0.16	9.25	9.88	-1.93	0.10	-2.14	-1.73
31340cac1	9.31	0.08	9.16	9.46	10.50	3.54	3.56	17.45	1.02	0.49	0.06	1.99
31340caf4	9.60	0.24	9.12	10.08	10.09	0.66	8.80	11.37	-1.38	0.25	-1.88	-0.88
31340cah0	9.07	0.12	8.85	9.30	8.95	0.53	7.91	9.98	-1.76	0.31	-2.36	-1.16
31340caj6	10.47	0.12	10.24	10.71	11.01	0.12	10.76	11.25	-0.66	0.13	-0.92	-0.40
31340can7	10.25	0.07	10.11	10.39	10.81	0.10	10.61	11.01	-1.17	0.44	-2.03	-0.32
31340cb55	11.06	0.20	10.68	11.45	13.41	1.18	11.09	15.73	-1.45	0.08	-1.60	-1.30
31340cba4	12.31	0.21	11.89	12.72	12.15	0.16	11.85	12.46	-0.82	0.17	-1.15	-0.49
31340cbg1	10.61	0.53	9.57	11.65	13.31	0.30	12.71	13.90	-1.23	0.10	-1.43	-1.03
31340cbj5	11.20	0.75	9.74	12.67	13.00	0.28	12.44	13.55	-1.13	0.08	-1.28	-0.98
31340cbk2	11.64	0.82	10.04	13.24	12.34	0.31	11.73	12.94	-1.18	0.09	-1.36	-0.99
31340cbl0	11.46	0.93	9.63	13.29	11.60	0.47	10.68	12.51	-0.97	0.14	-1.24	-0.70
31340cbp1	9.36	0.83	7.73	11.00	13.26	0.53	12.22	14.31	-1.38	0.12	-1.62	-1.15
31340cbr7	9.62	1.02	7.61	11.63	11.19	0.28	10.63	11.75	-1.12	0.12	-1.35	-0.89
31340cbs5	12.21	0.58	11.06	13.35	11.41	0.27	10.87	11.94	-1.27	0.15	-1.57	-0.97
31340cbw6	13.11	0.43	12.27	13.95	13.09	0.43	12.24	13.94	-1.25	0.13	-1.50	-1.01
31340ce60	9.17	0.56	8.08	10.26	9.30	0.98	7.38	11.22	-1.15	0.18	-1.49	-0.80
31340cej2	13.34	0.15	13.03	13.64	13.44	0.19	13.06	13.82	-1.14	0.20	-1.53	-0.76
31340cel7	9.14	0.11	8.92	9.35	3.03	3.84	-4.50	10.57	4.75	0.11	4.52	4.97
31340cer4	10.40	0.08	10.25	10.55	10.67	0.22	10.24	11.10	-1.26	0.13	-1.51	-1.01
31340cfr3	9.56	0.30	8.96	10.16	11.51	0.92	9.71	13.32	-2.72	0.72	-4.14	-1.30

Table IV
Estimated Values for the Parameters of the Bayesian Mixture of Regressions for Ginnie Mae Pools

Table IV shows the estimated values for model parameters, their standard deviation and 95% Probability intervals. Results are reported for all Ginnie Mae mortgage pools For example, for pool with CUSIP = 362026YJ9, the estimated value of the intercept is 9.39 with a standard deviation of 0.41. LPL stands for Left Probability Limit and RPL stands for Right Probability Limit.U_0 and V_0 are the intercepts, U_3 and V_3 are the burnout effect coefficients for the regressions chosen with probability p and 1-p, respectively.

	PARAMETER U_0				PARAMETER U_3				PARAMETER V_0				PARAMETER V_3			
Pool CUSIP	Average	SD	LPL	RPL	Average	SD	LPL	RPL	Average	SD	LPL	RPL	Average	SD	LPL	RPL
362026yj9	9.39	0.41	8.58	10.20	-0.73	0.45	-1.62	0.15	8.76	0.18	8.41	9.12	-0.71	0.06	-0.83	-0.60
3620274l5	9.65	0.10	9.46	9.83	-0.94	0.05	-1.04	-0.83	10.75	0.14	10.48	11.03	-0.53	0.08	-0.68	-0.37
3620274n1	9.74	0.02	9.70	9.77	-0.94	0.02	-0.97	-0.90	10.67	0.04	10.58	10.76	-0.75	0.04	-0.83	-0.68
3620274q4	10.11	0.05	10.01	10.20	-0.24	0.08	-0.39	-0.09	10.11	0.05	10.02	10.20	-1.52	0.05	-1.62	-1.43
3620276t6	11.14	0.25	10.64	11.63	-0.67	0.16	-0.99	-0.35	10.09	0.16	9.77	10.41	-0.97	0.08	-1.13	-0.81
362027bx1	11.38	0.04	11.30	11.46	-0.76	0.03	-0.82	-0.70	10.85	0.03	10.78	10.91	-0.93	0.04	-1.00	-0.85
362027fq2	9.85	0.11	9.63	10.08	-0.69	0.35	-1.37	-0.01	10.46	0.29	9.90	11.02	-1.14	0.29	-1.72	-0.57
362027hs6	9.72	0.22	9.29	10.15	0.28	1.32	-2.30	2.87	10.09	0.57	8.98	11.20	-1.20	0.34	-1.86	-0.54
362027hz0	10.81	0.06	10.70	10.93	-0.58	0.05	-0.68	-0.49	9.61	0.03	9.56	9.66	-0.92	0.02	-0.95	-0.88
362027j47	10.63	0.04	10.55	10.70	-0.87	0.02	-0.90	-0.83	9.93	0.04	9.85	10.01	-1.11	0.02	-1.15	-1.06
362027jp0	10.46	0.04	10.39	10.54	-0.67	0.04	-0.76	-0.58	9.78	0.02	9.73	9.83	-0.78	0.02	-0.82	-0.75
362027kb9	10.49	0.03	10.44	10.55	-0.77	0.03	-0.83	-0.72	9.64	0.01	9.61	9.67	-0.76	0.02	-0.79	-0.73
362027ke3	8.60	0.11	8.39	8.81	-0.75	1.84	-4.36	2.85	10.01	0.67	8.69	11.33	-3.10	0.91	-4.88	-1.31
362027kw3	9.61	0.02	9.58	9.65	-0.93	0.02	-0.96	-0.89	10.74	0.03	10.67	10.80	-0.99	0.02	-1.03	-0.94
362027n26	11.19	0.04	11.12	11.26	-0.88	0.03	-0.95	-0.82	10.35	0.03	10.30	10.40	-1.03	0.02	-1.06	-1.00
362027qy3	9.77	0.16	9.46	10.08	-0.26	0.54	-1.31	0.80	9.98	0.31	9.38	10.57	-1.30	0.28	-1.84	-0.76
362027r55	9.89	0.09	9.72	10.06	-0.73	0.25	-1.21	-0.25	8.69	0.41	7.88	9.49	-0.38	0.26	-0.89	0.13
362027rc0	11.22	0.03	11.17	11.27	-0.75	0.02	-0.80	-0.71	10.50	0.02	10.46	10.54	-0.81	0.01	-0.84	-0.78
362027ss4	10.77	0.02	10.73	10.81	-0.61	0.04	-0.68	-0.54	9.96	0.02	9.92	10.00	-0.63	0.01	-0.66	-0.60
362027t95	11.58	0.03	11.52	11.63	-0.93	0.05	-1.02	-0.84	11.06	0.02	11.02	11.10	-0.98	0.02	-1.02	-0.94
362027uj1	10.91	0.03	10.84	10.97	-0.90	0.04	-0.98	-0.83	10.49	0.02	10.44	10.53	-1.45	0.04	-1.53	-1.37
362027uk8	11.22	0.16	10.91	11.52	-0.84	0.09	-1.02	-0.65	10.50	0.05	10.40	10.59	-0.93	0.03	-0.98	-0.87
362027v76	10.74	0.05	10.65	10.84	-0.70	0.04	-0.77	-0.63	10.11	0.02	10.07	10.15	-0.89	0.01	-0.90	-0.87
362028af9	9.54	0.11	9.32	9.76	1.40	1.16	-0.88	3.68	10.85	0.24	10.38	11.31	-7.79	1.69	-11.10	-4.48
362028az5	11.30	0.16	10.98	11.62	-0.72	0.34	-1.39	-0.06	10.64	0.10	10.45	10.83	-1.02	0.09	-1.19	-0.85
362028bt8	10.72	0.17	10.37	11.06	-0.59	0.10	-0.79	-0.39	9.66	0.09	9.48	9.85	-0.97	0.03	-1.04	-0.91
362028bx9	10.99	0.09	10.81	11.16	-0.65	0.08	-0.80	-0.49	9.64	0.02	9.60	9.68	-0.96	0.02	-1.00	-0.92
362028dy5	10.72	0.11	10.50	10.93	-0.35	0.49	-1.32	0.62	11.32	0.16	11.02	11.63	-1.46	0.12	-1.70	-1.22
362028eq1	12.19	0.03	12.13	12.24	-1.18	0.05	-1.29	-1.08	11.49	0.02	11.45	11.52	-1.07	0.02	-1.11	-1.04
362028f31	9.79	0.37	9.07	10.51	-0.39	0.60	-1.57	0.79	10.18	0.38	9.43	10.92	-1.25	0.15	-1.55	-0.96
362028fb3	10.64	0.08	10.48	10.79	-0.26	0.20	-0.66	0.13	11.27	0.12	11.03	11.51	-1.53	0.06	-1.66	-1.41
362028fm9	11.32	0.12	11.08	11.57	-0.71	0.22	-1.14	-0.29	10.60	0.07	10.46	10.74	-0.95	0.03	-1.01	-0.90
362028hb1	11.02	0.05	10.92	11.13	-0.88	0.04	-0.97	-0.80	9.55	0.02	9.52	9.58	-0.76	0.01	-0.79	-0.74
362028hu9	9.63	0.03	9.58	9.69	-0.95	0.02	-0.98	-0.92	10.65	0.04	10.57	10.72	-0.50	0.03	-0.56	-0.44
362028jf0	10.81	0.05	10.71	10.92	-0.49	0.05	-0.59	-0.39	9.87	0.04	9.79	9.94	-0.92	0.02	-0.97	-0.88
362028jk9	10.47	0.03	10.41	10.52	-0.65	0.06	-0.78	-0.53	11.33	0.04	11.25	11.41	-1.34	0.02	-1.38	-1.31

Table V
Summary of the Estimated Coeffcients Across all Pools

Table V summarizes the estimation results across all pools. Panel A gives the summary for Ginnie Mae pools and Panel B gives the summary for Freddie Mac pools. U_0, V_0 are the intercepts, U_3 and V_3 are the burnout coefficients.

Panel A: Ginnie Mae Averages Across Pools

PARAMETER U_0			PARAMETER U_3			PARAMETER V_0			PARAMETER V_3		
Average	LPL	RPL	Average	LPL	RPL	Average	LPL	RPL	Average	LPL	RPL
10.50	10.31	10.69	-0.62	-1.11	-0.13	10.28	10.04	10.53	-1.23	-1.50	-0.96

Panel B: Freddie Mac Averages Across Pools

PARAMETER U_0			PARAMETER V_0			PARAMETER V_3		
Average	LPL	RPL	Average	LPL	RPL	Average	LPL	RPL
10.27	9.69	10.84	10.58	9.39	11.77	-0.94	-1.38	-0.50

Figure 1

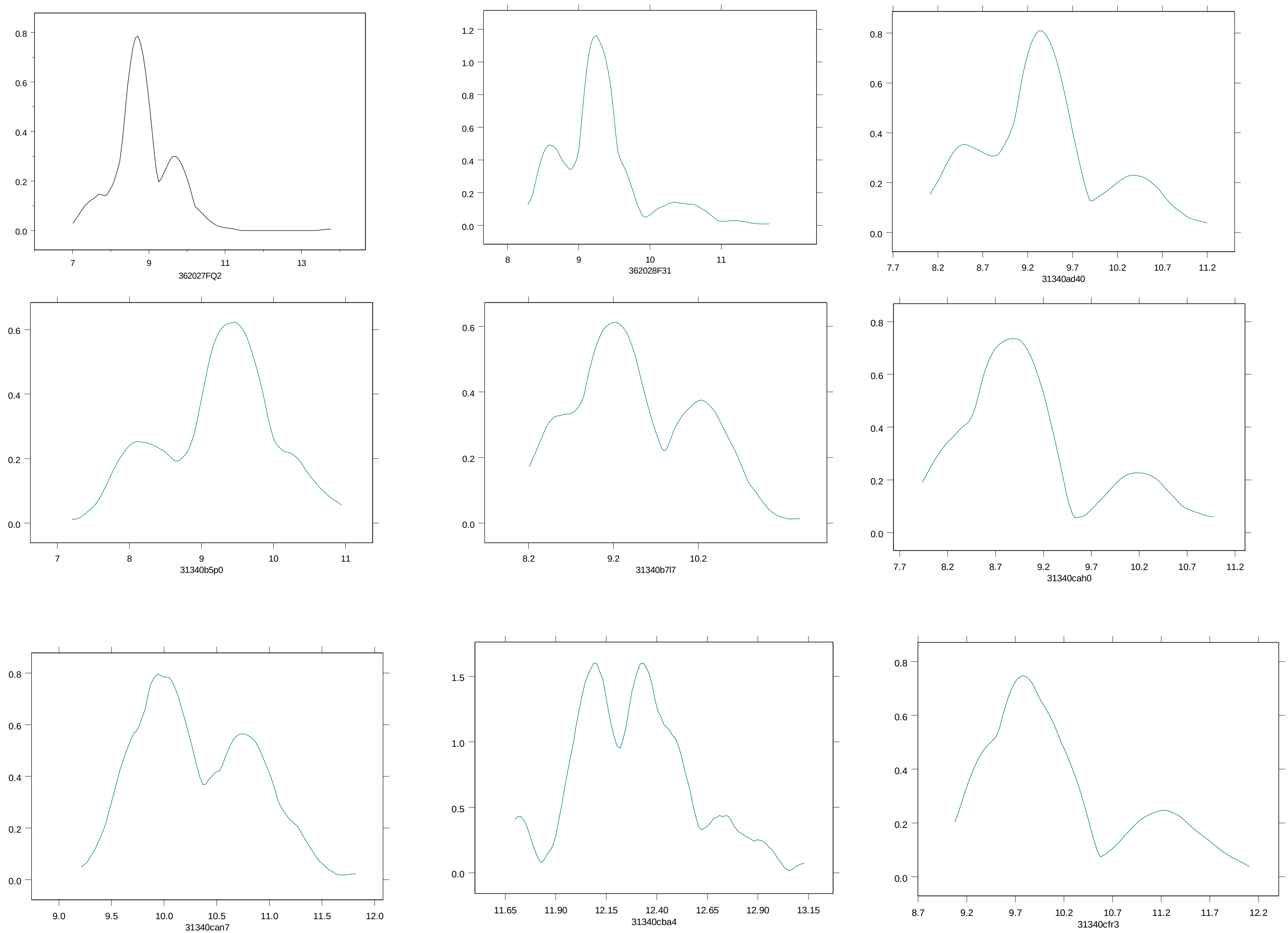


Figure 1 shows densities of the logarithm of the actual dollar amount paid for nine pools of mortgages.

Figure 2

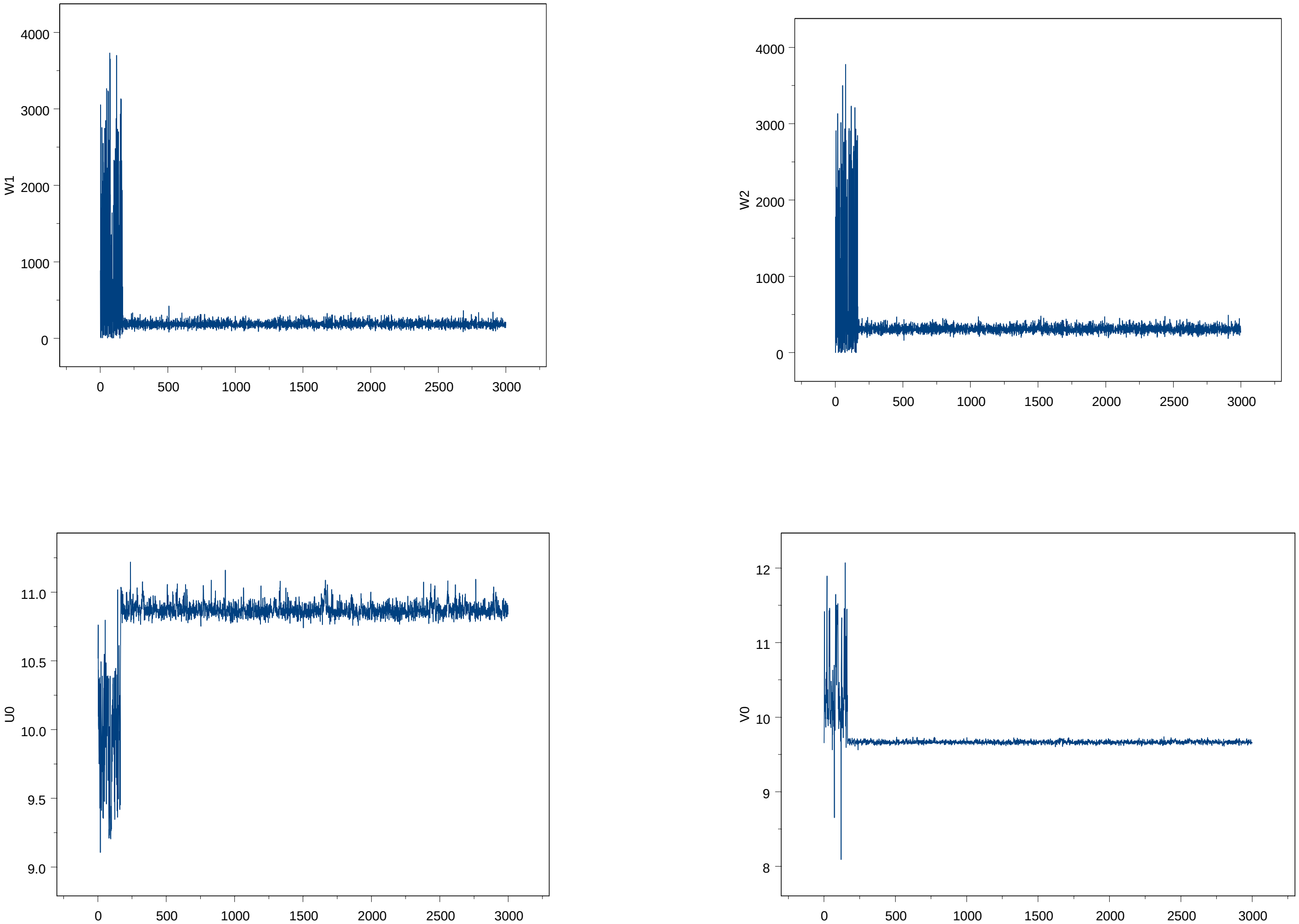


Figure 2 shows the actual simulated points for U_0, V_0, W_1 and W_2 for one of the pools of mortgages

Figure 3

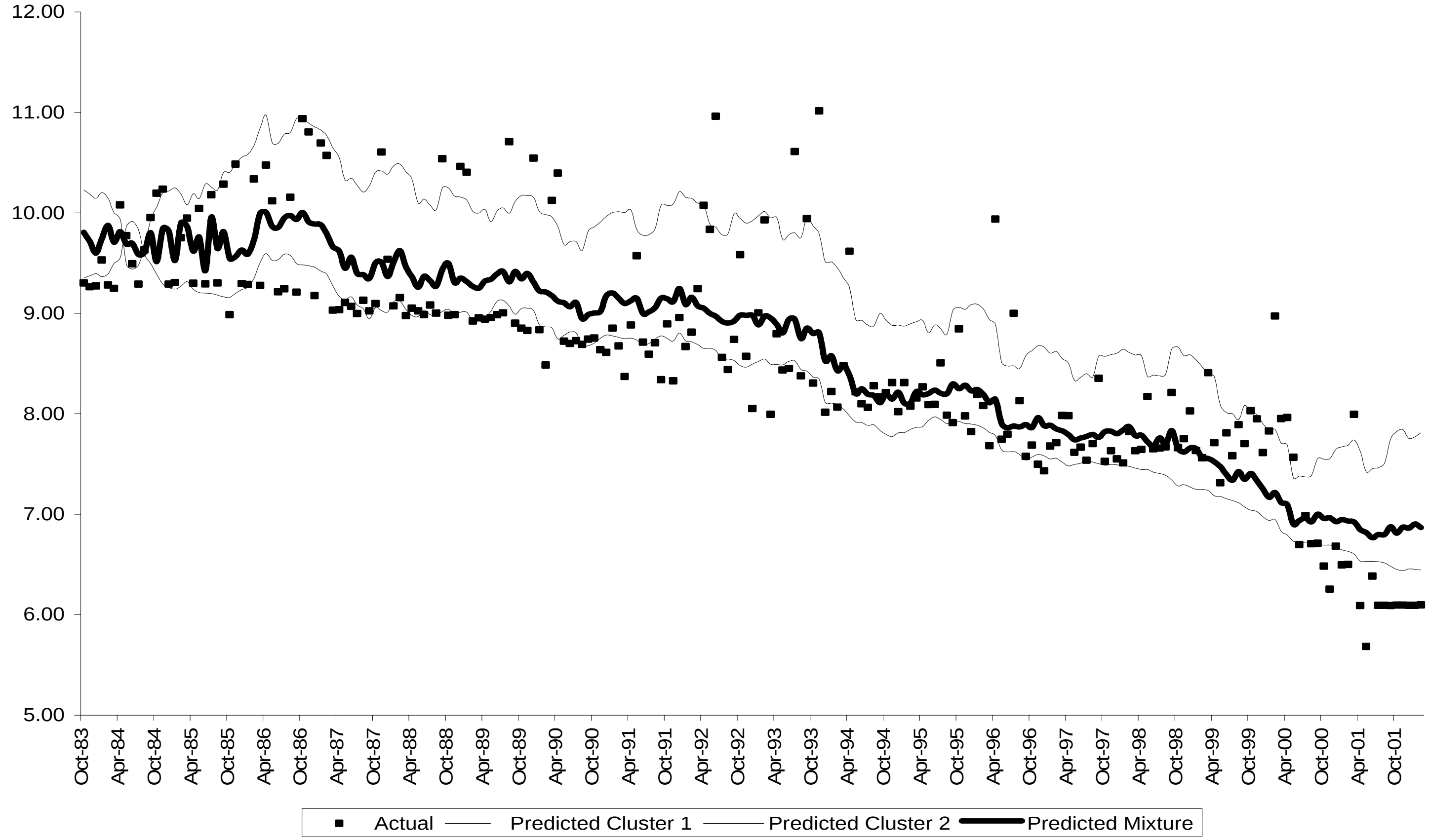


Figure 3 compares the actual prepayment with the predicted prepayment for one pool of mortgages, CUSIP 31340ABB6.

The lower solid line is $\sum_{k=0}^{5} \hat{U}_k X_k(t)$, where $\hat{U}_k$ are the expected values of the parameters posterior distributions.

The upper solid line is $\sum_{k=0}^{5} \hat{V}_k X_k(t)$, where $\hat{V}_k$ are the expected values of the parameters posterior distributions.

A point forecast is the weighted average of these two lines (the middle solid line).

Figure 4

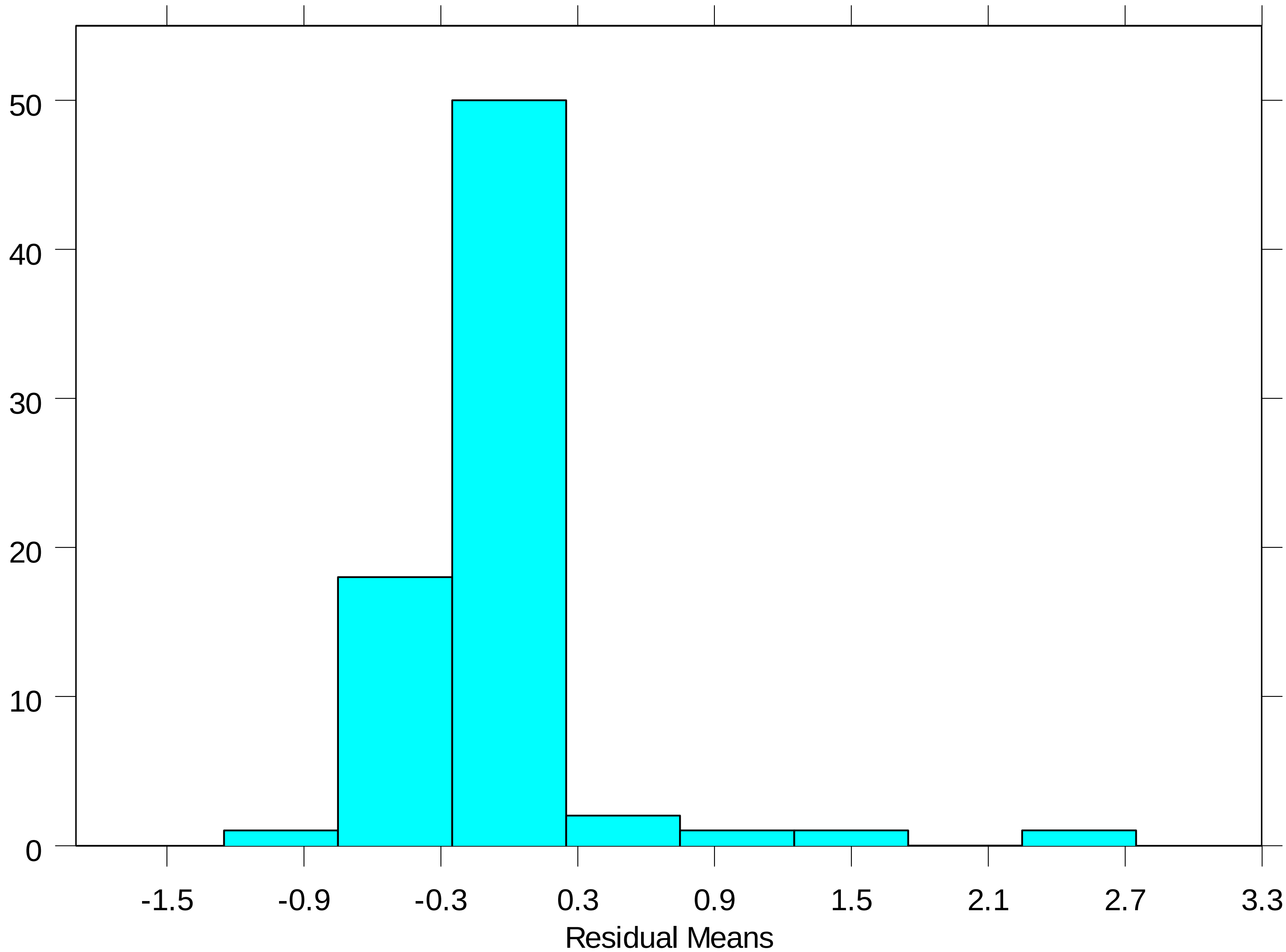
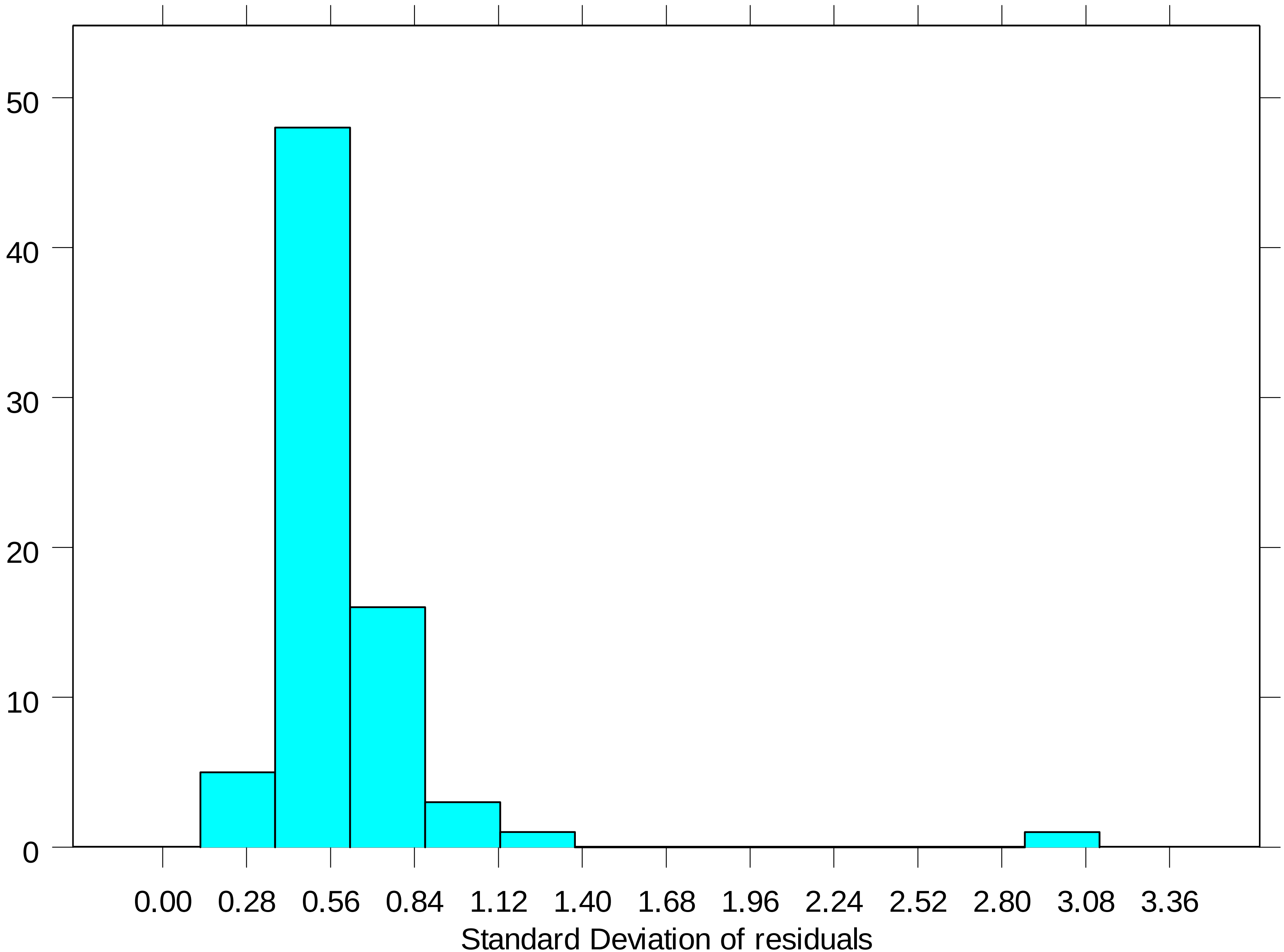


Figure 5



50
40
30
20
10
0
0.00 0.28 0.56 0.84 1.12 1.40 1.68 1.96 2.24 2.52 2.80 3.08 3.36
Standard Deviation of residuals