Title
Pool Heterogeneity and the Valuation of Mortgage-Backed Securities

Permalink
https://escholarship.org/uc/item/74b062hm

Author
Stanton, Richard

Publication Date
1992-04-01

Peer reviewed
CENTER FOR REAL ESTATE AND URBAN ECONOMICS

WORKING PAPER SERIES

WORKING PAPER NO. 92-208

POOL HETEROGENEITY AND THE VALUATION OF MORTGAGE-BACKED SECURITIES

By

RICHARD STANTON

These papers are preliminary in nature: their purpose is to stimulate discussion and comment. Therefore, they are not to be cited or quoted in any publication without the express permission of the author.

WALTER A. HAAS SCHOOL OF BUSINESS
CENTER FOR REAL ESTATE AND URBAN ECONOMICS
UNIVERSITY OF CALIFORNIA AT BERKELEY
Kenneth Rosen, Chair

The Center was established in 1950 to examine in depth a series of major changes and issues involving urban land and real estate markets. The Center is supported by both private contributions from industry sources and by appropriations allocated from the Real Estate Education and Research Fund of the State of California.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH
Richard Sutch, Director

The Institute of Business and Economic Research is an organized research unit of the University of California at Berkeley. It exists to promote research in business and economics by University faculty. These working papers are issued to disseminate research results to other scholars. The authors welcome comments; inquiries may be directed to the author in care of the Center.
Pool Heterogeneity and the Valuation of Mortgage-Backed Securities.

Richard Stanton

Haas School of Business
U.C. Berkeley

First Draft: June 1991
This Draft: April 1992

WORKING PAPER #92-208

---

This paper is based on a chapter from my Ph.D. dissertation. I am grateful for helpful comments and suggestions to Darrell Duffie, Robert Whitelaw and Nicky King. I especially wish to thank the members of my dissertation committee, Allan Kleidon, Ken Singleton and Peter Reiss, for their enormous contributions. I thank Walter Torous and a Wall Street firm for providing the data used in this study, and I gratefully acknowledge financial assistance from the Berkeley Program in Finance, from the Dean Witter Foundation, and from the U.C. Berkeley Center for Real Estate and Urban Economics.
ABSTRACT

Stanton [38] developed a structural model of mortgage prepayment that explained mortgage prepayment better than the descriptive model of Schwartz and Torous [36], based on a comparison of the fraction of variation explained by each model. However, while that model allows heterogeneity in the transaction costs faced by the individual mortgage holders within a pool, each pool is assumed to be identical apart from size. The model does not allow for heterogeneity between mortgage pools.

This paper extends the model of rational mortgage prepayment developed in Stanton [38] to allow for unobservable heterogeneity in the distribution of transaction costs between different mortgage pools. One major source of such heterogeneity is the different geographical location of mortgages in separate pools.

Unobservable heterogeneity among mortgage pools raises problems in pricing mortgage-backed securities. In order to price the securities properly, account must be taken of this heterogeneity. As time goes by, more information about the pool's type arrives in the form of its prepayment history. This information should be used to update beliefs about the pool, future prepayments, and hence security value. Traders on Wall Street recognize that this heterogeneity exists. They talk about “fast pay” and “slow pay” pools. However, previous methods for generating pool-specific security prices are ad hoc. This paper presents a consistent methodology for calculating conditional expectations of future prepayment levels, and hence for deriving pool-specific security prices that incorporate all available prepayment information. Knowledge of pool-specific prices is important both for valuation purposes and for determining the cheapest pool to deliver when selling mortgage-backed securities.

The new model is estimated using non-linear least squares on the same dataset used to estimate the model in Stanton [38]. It explains observed prepayment behavior significantly better than the basic model. The model’s prediction errors are regressed against several economic variables that are often used to predict mortgage prepayment. Few of the variables appear significantly in the regression, and correcting for the results of the regression makes little difference to the predicted prepayment levels.
1 Introduction

Stanton [38] developed a structural model of mortgage prepayment that explained mortgage prepayment better than the descriptive model of Schwartz and Torous [36], based on a comparison of the fraction of variation explained by each model. However, while that model allows heterogeneity in the transaction costs faced by the individual mortgage holders within a pool, each pool is assumed to be identical apart from size. The model does not allow for heterogeneity in the distribution of transaction costs between mortgage pools.\footnote{Factors which might induce such heterogeneity include the different geographical location of the mortgages in different pools, legal factors (for example, whether a mortgage-holder can just walk away from his or her mortgage depends on the state involved), different education levels, different propensities to change jobs and relocate etc.} Because of this, information about an individual pool does not affect beliefs about future prepayment and security value. Given unobservable heterogeneity among pools, the monthly prepayment rate for each pool provides information that allows us to update beliefs about its type. This allows expectations of future prepayment to be calculated conditional on all available information, and these expectations can be used to generate pool-specific mortgage-backed security prices month after month. Knowledge of pool-specific prices is important both for valuation purposes and for determining the cheapest pool to deliver when selling mortgage-backed securities. Traders on Wall Street recognize that this heterogeneity exists. They talk about “fast pay” and “slow pay” pools. However, previous methods for generating pool-specific security prices are ad hoc. One method involves multiplying the generic premium above par by the ratio of the fraction of the pool’s principal remaining to the fraction for an “average” pool. While this adjustment is simple to apply, it is not based on any specific model of heterogeneity, so there is no particular reason to believe it gives correct results.

This paper extends the prepayment model of Stanton [38] to incorporate unobservable heterogeneity in the distribution of transaction costs between different mortgage pools. It presents a consistent methodology for calculating conditional expectations of future prepayment levels, and hence for deriving pool-specific security prices that incorporate all available information. The new model is estimated using non-linear least squares on the same dataset used to estimate the model in Stanton [38]. The heterogeneous model explains observed prepayment behavior significantly better than the basic model. The model’s prediction errors are regressed against several economic variables that are often used to predict mortgage
prepayment. Few of the variables appear significantly in the regression, and correcting for the results of the regression makes little difference to the predicted prepayment levels.

The paper is organized as follows. Sections 2 and 3 lay out the model and describe its detailed implementation. The model in Stanton [38] is summarized, and the several extensions necessary to handle heterogeneous mortgage pools are detailed. Section 4 estimates the extended model based on its predictions for observed prepayment behavior. The same dataset is used as in Stanton [38], though non-linear least squares is used to perform the estimation, rather than generalized method of moments. Using non-linear least squares gives enough degrees of freedom to estimate the number of parameters required to describe fully the heterogeneity that exists among pools.

Section 6 discusses how to value mortgage-backed securities based on the prepayment model previously described. It lays out in detail the updating process that incorporates fully the information contained in the history of monthly prepayment rates, and uses this to derive pool-specific security prices. Section 7 discusses extensions of this work, including the possibility of using prices generated by the model as another means of estimating the model by comparing these prices with market prices, if available. Concluding remarks are given in section 8.

2 Structure of the Model

The model described in this paper is an extension of that in Stanton [38]. Mortgage prepayment is either the result of an interest related option exercise decision or is for some exogenous reason, such as forced relocation. Each mortgage-holder faces a transaction cost payable on prepayment. Exogenous prepayment is governed by a hazard rate $\lambda$. The probability of prepayment for exogenous reasons in a time interval of length $\delta t$ starting at $t$, conditional on not having prepaid prior to $t$, is approximately $\lambda \delta t$. The mortgage holder decides whether to prepay the mortgage at random intervals, governed by the constant parameter $\rho$, which defines an exponential distribution describing the time between successive decision points. Similar to the process governing exogenous prepayments, the probability that the next decision is made in a time interval of length $\delta t$ starting at $t$, conditional on not having made the next decision prior to $t$, is approximately $\rho \delta t$.

Given known values of $\lambda, \rho$ and the transaction cost payable on prepayment, the mortgage-
holder chooses a prepayment strategy which maximizes the market value of his prepayment option, or equivalently minimizes the market value of his mortgage liability.

2.1 Heterogeneity

Heterogeneity among the mortgage-holders within a pool is modeled as in Stanton [38] by assuming that different mortgage-holders face different transaction costs, initially distributed according to a beta distribution. However, not only do individual pools contain mortgages held by mortgage-holders of different types, but different pools may themselves have very different characteristics. A major reason often cited for this heterogeneity is geographical.\(^3\) This occurs for a number of reasons, including geographical variation in education levels, local economic conditions and legal environments. If we knew where mortgages in a pool were concentrated, we could adjust for this by including location variables in the model. However, investors do not usually have this information. They may know the servicing institution, but this is often a national institution, and may not be the original issuer. Differences between pools may exist even after controlling for the state of issue. Education levels will vary widely, which may have a significant impact on how ready mortgage-holders are to prepay should it be optimal to do so. Similarly, mortgage-holders in two pools may have very different propensities to relocate for job or other reasons. A number of factors therefore lead to unobservable heterogeneity between pools. These same factors presumably also explain much of the heterogeneity within pools discussed in Stanton [38].

Heterogeneity between pools is modeled here by allowing the initial distribution of transaction costs to vary across pools. An obvious way to achieve this would be to allow the mean of the cost distribution to vary across pools, while keeping the variance fixed. However, there are drawbacks to this approach. The mean and variance of a beta distribution are

\[
\mu = \frac{\alpha}{\alpha + \beta}, \\
\sigma^2 = \frac{\alpha \beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}.
\]

The inverse, giving \(\alpha\) and \(\beta\) in terms of \(\mu\) and \(\sigma^2\), is

\[
\alpha = \mu \left[ \frac{\mu(1 - \mu)}{\sigma^2} - 1 \right],
\]

\(^3\)Beckett and Morris [3] present evidence that prepayment behavior varies by state.
\[ \beta = (1 - \mu) \left[ \frac{\mu(1 - \mu)}{\sigma^2} - 1 \right]. \]

Both \( \alpha \) and \( \beta \) must be positive for a valid beta distribution, so the expression in brackets must be positive. In other words,

\[ \mu(1 - \mu) > \sigma^2. \]  \( \text{(1)} \)

For a given value of \( \sigma^2 \), this imposes limits on possible values for \( \mu \). Assuming \( \sigma^2 < 1/4 \), \( \mu \) may only take on values between the two (positive) roots of the equation

\[ \mu^2 - \mu + \sigma^2 = 0. \]  \( \text{(2)} \)

These roots are

\[ \mu = \frac{1}{2} \pm \frac{1}{2} \sqrt{1 - 4\sigma^2}. \]  \( \text{(3)} \)

Thus \( \mu \) may not be very close to either 0 or 1. As \( \sigma^2 \) approaches 1/4, the permissible range becomes narrower. For \( \sigma^2 = 1/4 \), the only solution to the equation is \( \mu = 1/2 \). For \( \sigma^2 > 1/4 \) the equation has no real solutions. To allow for maximum flexibility in the mean of the cost distribution, we shall instead assume that all pools share the same value of the parameter \( \alpha \), but pool \( i \) has a parameter \( \beta_i \) which is specific to that pool. By choosing values of \( \beta \) between 0 and \( \infty \), the mean of the cost distribution varies between 0 and 1, the full range of possible values. The partial derivatives of \( \mu \) and \( \sigma^2 \) with respect to \( \beta \) are

\[ \frac{\partial \mu}{\partial \beta} = \frac{-\alpha}{(\alpha + \beta)^2}, \]  \( \text{(4)} \)

\[ \frac{\partial \sigma^2}{\partial \beta} = \frac{\alpha (\alpha^2 + \alpha - \alpha \beta - \beta - 2\beta^2)}{(\alpha + \beta)^3 (\alpha + \beta + 1)^2}. \]  \( \text{(5)} \)

The mean of the transaction cost distribution decreases with \( \beta \). The variance increases with \( \beta \) if the expression in brackets in the numerator of equation 5 is positive, and decreases if it is negative. There is a single positive value of \( \beta \) at which the expression is zero,

\[ \beta_0(\alpha) = \frac{1}{4} \left[ \sqrt{9\alpha^2 + 10\alpha + 1} - (\alpha + 1) \right]. \]  \( \text{(6)} \)

For \( \beta > \beta_0(\alpha) \) the expression is negative, and the variance decreases in \( \beta \). For \( \beta < \beta_0(\alpha) \) the expression is positive, and the variance increases in \( \beta \). As \( \beta \to \infty \), \( \mu \to 0 \) and \( \sigma^2 \to 0 \). As \( \beta \to 0 \), \( \mu \to 1 \) and \( \sigma^2 \to 0 \).

There are other possible specifications of heterogeneity between pools that could have been used. For example, instead of allowing the transaction cost distribution to vary between
pools, the exogenous prepayment parameter $\lambda$ or the parameter governing the time between prepayment decisions, $\rho$, could have been allowed to vary. However, either of these alternative specifications would have resulted in a model which was computationally infeasible. Modeling between pool heterogeneity via the parameters of the transaction cost distribution ensures that the model remains feasible to implement and estimate.\textsuperscript{4}

3 Implementation

As in Stanton [38], we shall use the Cox, Ingersoll and Ross one factor model, with parameter values estimated by Pearson and Sun [33].

\[
\begin{align*}
\kappa &= 0.29368, \\
\mu &= 0.07935, \\
\sigma &= 0.11425, \\
q &= -0.12165.
\end{align*}
\]

To value a single mortgage liability, simultaneously determining the optimal prepayment strategy for the mortgage, the Crank-Nicholson finite difference algorithm is used to solve the partial differential equation for asset prices,

\[
\frac{1}{2} \sigma^2 r V_{rr} + [\kappa \mu - (\kappa + q) r] V_r + V_t - r V + C = 0, \tag{7}
\]

on a grid of time and interest rate values, subject to appropriate boundary conditions. In Stanton [38] the continuous distribution of transaction cost values in a pool was approximated by a distribution taking on only a finite number of values. Repeating the valuation procedure for mortgages of each possible prepayment cost level allowed us to determine the critical cost level for each time and interest rate level, the level above which it is optimal not to prepay and below which it is optimal to prepay. This, combined with the proportions of each cost type in a pool, gave the proportion of the pool prepaying at each time. The value of a mortgage-backed security backed by each mortgage type was calculated at the same time, using the exercise strategy determined for the mortgage-holder to determine the cash flows accruing to the security.

\textsuperscript{4}See section 3.
3.1 Heterogeneity between Pools

Consider now implementing the extended model with a separate initial transaction cost distribution for each mortgage pool. For each pool, we could proceed as described for the basic model, calculating a critical transaction cost grid, and hence expected prepayment levels for each period. However, this would need to be repeated for each pool, as a new grid would be generated for each set of values \((\alpha, \beta, \lambda, \rho)\). To perform the analysis for a large number of mortgage pools would quickly become too time consuming to be practical.

If the heterogeneity were modeled by allowing one of the parameters \(\lambda\) or \(\rho\) to vary across pools, this would present a problem. Modifying one of these parameters changes the whole grid of values and the optimal exercise strategy for mortgage holders with every level of prepayment cost. Thus, every time a new value of \(\lambda\) or \(\rho\) is considered, a new grid has to be generated. This would take a prohibitive amount of time. However, when the heterogeneity is over one of the cost parameters, there is a way around the problem. Note that the values in the critical cost grid depend on the values of \(\lambda\) and \(\rho\) and the discrete transaction costs used to form the discrete approximation. They do not depend explicitly on the actual values of \(\alpha\) and \(\beta\). The dependence on these parameters is only through the way these parameters are used to calculate the transaction cost values used in the discrete approximation.

Modifying the discrete approximation used in Stanton [38] slightly allows implementation of the heterogeneous model for a single set of \(\rho, \lambda\) values with only a single critical cost grid calculation. Calculate a single set of approximation points \(X_1, X_2, \ldots, X_m\), as in Stanton [38], for a single \(\beta\) value, say the arithmetic mean of the values for each pool. Now calculate a critical transaction cost grid corresponding to those cost values. For each pool, let the initial transaction cost distribution be approximated by the same set of points, \(X_1, X_2, \ldots, X_m\). Let \(B_i(x)\) be the cumulative distribution function for the initial cost distribution of pool \(i\). Then let the weight corresponding to cost level \(X_j\) be

\[
    c_{ij} = \begin{cases} 
    \frac{B_i(X_j) + B_i(X_2)}{2} & \text{if } j = 1, \\
    \frac{B_i(X_{j+1}) - B_i(X_{j-1})}{2} & \text{if } 2 < j < m, \\
    1 - \frac{B_i(X_{m-1}) + B_i(X_m)}{2} & \text{if } j = m.
    \end{cases}
\]  

By keeping the points the same, and varying the associated weights, we can use a single grid to handle multiple pools. To see how this new approximation looks in practice, figures 1 and 2 show the approximation for two sets of parameter values. The discrete points for both
Figure 1: Discrete approximation to beta distribution. Approximation points calculated using parameter values $\alpha = 3.0, \beta = 4.0$. Weights calculated using the same parameter values.

Approximations are calculated as in Stanton [38] using parameters $\alpha = 3.0, \beta = 4.0$. Figure 1 shows the discrete approximation calculated from equation 8 for the same parameter values $\alpha = 3.0, \beta = 4.0$. This illustrates the fact that if $\beta_j = \beta$, the value used to calculate the grid, the weight on each point is $1/m$, so we have the same approximation as before. Figure 2 shows the discrete approximation calculated using the same points, but parameter values $\alpha = 3.0, \beta = 2.0$. The discrete c.d.f. generated by equation 8 has the property (shared with the approximation used in Stanton [38]) that at each discrete cost value, it jumps to the value halfway between the true c.d.f. at that value and the true c.d.f. at the next higher discrete cost value (subject to starting at 0 and ending at 1). Thus the amount that the discrete approximation overshoots the true c.d.f. as the cost level approaches $X_j$ from above exactly equals the amount by which the true c.d.f. exceeds the discrete approximation as the cost level approaches $X_{j+1}$ from below. This approximation allows us to use only a single grid but to approximate the cost distribution for a range of values of $\beta_i$. This immensely reduces the amount of computation time required, and makes the model feasible to implement and estimate.
Figure 2: Discrete approximation to beta distribution. Approximation points calculated using parameter values $\alpha = 3.0, \beta = 4.0$. Weights calculated using different parameter values $\alpha = 3.0, \beta = 2.0$. 
4 Estimation

In Stanton [38] the basic prepayment model was estimated using a version of Hansen’s generalized method of moments (GMM). This is a particular example of an estimation technique where, writing $w_{it}$ for the observed proportion of pool $i$ prepaying in month $t$, and $\bar{w}_t$ for the expected value of this proportion conditional on some or all of the information available at time $t$, parameter values are picked to minimize a function of the difference between them. In this paper, we specialize the methodology to allow the estimation of a large number of parameters. Instead of taking as moments the average prediction error over all pools for each month, we shall use each monthly prediction error for each pool as a separate moment, resulting in an effective sample size of 1. Since this is not sufficient to permit the calculation of the usual second stage GMM weighting matrix, we need to impose restrictions on the variance matrix of the residuals. Taking this to be a scalar multiple of the identity matrix makes this estimation procedure exactly equivalent to non-linear least squares (NLLS). If

$$y_t = f_t(\theta_o) + u_t, \quad t = 1, 2, \ldots, T,$$  \hfill (9)

where $y_t$ is an observable random variables, $\theta_o$ is a $K$-vector of (unknown) parameters, and \{u_t\} are i.i.d. with $E(u_t) = 0$ and $\text{var}(u_t) = \sigma_o^2$ for all $t$, then $\hat{\theta}$, the NLLS estimator of $\theta_o$, is the value which minimizes the sum of squared residuals

$$S_T(\theta) = \sum_{t=1}^{T} [y_t - f_t(\theta)]^2.$$ \hfill (10)

The NLLS estimator of $\sigma_o^2$ is

$$\hat{\sigma}^2 = \frac{S_T(\hat{\theta})}{T}.$$ \hfill (11)

The asymptotic variance-covariance matrix for the estimated parameter values is

$$\frac{S_T(\hat{\theta})}{T} \left[ \sum_{t=1}^{T} \frac{\partial f_t}{\partial \theta} \frac{\partial f_t}{\partial \theta'} \right]^{-1}.$$ \hfill (12)

4.1 Data

As in Stanton [38] the prepayment data used for estimation were monthly prepayment rates for 12.5% GNMA 30 year single family mortgages over the period July 1983 – December 1989. Pools with missing data were excluded, leaving 1,156 pools in the sample used for estimation. The Salomon Brothers yield on newly issued 20 year Treasury bonds was used
to derive a short rate series to feed into the CIR model as in Stanton [38]. For graphs, summary statistics and further discussion of the process used to generate an implied short term interest rate series, see Stanton [38].

4.2 Incorporating Heterogeneity

The extended model, with a separate $\beta$ parameter for each pool, can be written in the form

$$w_{it} = \bar{w}_{it}(\rho, \lambda, \alpha, \beta_i) + v_{it},$$  \hspace{1cm} (13)

where $i = 1, 2, \ldots, 1156$ and $t = 1, 2, \ldots, 78$. Write $y$ for the vector formed by stacking the 1156 vectors $w_i$ on top of each other. Write $u$ for the vector formed similarly from the matrix $v_{it}$. The new series are defined by

$$y_{78(i-1)+t} = w_{it},$$  \hspace{1cm} (14)

$$u_{78(i-1)+t} = v_{it}. $$  \hspace{1cm} (15)

All observations for a single pool appear in time order, followed by all residuals for the next pool. Thus $y_1, y_2, \ldots, y_{78}$ are the 78 monthly prepayment rates for pool 1; $y_{79}, y_{80}, \ldots, y_{156}$ are the 78 monthly prepayment rates for pool 2, and so on. The model can now be written as

$$y_k = f_k(\theta) + u_k,$$  \hspace{1cm} (16)

where $\theta = (\rho, \lambda, \alpha, \beta_1, \beta_2, \ldots, \beta_{1156})$. Note that the index $k$ does not just run over time, or over pools, but runs from 1 to $78 \times 1156 = 90,168$, covering all pools and all time periods. We can regard the $\beta_i$ as nonlinear fixed effects.\footnote{An alternative approach would be to regard the $\beta_i$ as nonlinear random effects. In this case, each $\beta_i$ would be treated as a draw from a particular distribution. Knowledge about this distribution could be used to help estimate the parameters. However, since we know nothing a priori about the distribution of the $\beta_i$, we shall use the fixed effects model. For further discussion, within the framework of a linear model, see Judge et al. [28].} In the linear case each $\beta_i$ would force the mean of the residuals for pool $i$ to be zero, though since $f_k$ is a nonlinear function of $\beta_i$, this exact condition no longer holds. Since the number of parameters being estimated increases linearly with the number of mortgage pools, the asymptotic results we rely on apply as $T \to \infty$, not as $N \to \infty$. We have to estimate a total of 1,159 separate parameters. Given the amount of time required to estimate the basic model, with only 4 parameters, this seems to be an almost impossible task. However, one feature of the model simplifies the task of estimation...
enormously, and makes it of approximately the same order of complexity as estimating the basic model without heterogeneity. Looking at equation 13, we see that the residuals for pool \( i \) depend only on the value of \( \beta_i \), and not on the value of \( \beta \) for any other pool. Thus we do not need to consider the interaction between changes in different \( \beta \) values. Changing a particular \( \beta \) will only affect the residuals for a single pool, and the size and direction of this change is independent of the values of all other \( \beta \) parameters (though not of \( \rho, \lambda, \alpha, \) in general). For this to be true, we require there to be no correlation across pools, since correcting for this would make every residual depend on every \( \beta_i \). Such correlation does not exist under the null hypothesis that the model is correct, but may exist in practice if the model systematically misses a significant predictor of mortgage prepayment. Section 5 estimates this correlation, and finds it to be small. Write the sum of squares as \( S_K(\theta^*, \beta) \), where

\[
\theta^* = (\rho, \lambda, \alpha), \\
\beta = (\beta_1, \beta_2, \ldots, \beta_{1156}).
\]

Rather than minimizing \( S_K \) over 1,159 parameters simultaneously, we can minimize it in two steps. Define

\[
S^*(\theta^*) = S_K(\theta^*, \hat{\beta}(\theta^*)), \tag{17}
\]

where \( \hat{\beta}(\theta^*) \) is the set of \( \beta_i \) values which minimizes the sum of squared residuals conditional on the given values of \( \rho, \lambda, \alpha \). Choose \( \hat{\theta}^* \) to minimize the concentrated sum of squares \( S^*(\theta^*) \).

The independence of the residuals for each pool from the \( \beta \) value for all other pools implies that

\[
\min_{\beta} S_K(\theta^*, \beta) = \sum_{i=1}^{I} \min_{\beta_i} S_i(\theta^*, \beta_i), \tag{18}
\]

where \( S_i(\theta^*, \beta_i) \) is the sum of squares for pool \( i \) only. Thus, in calculating the concentrated sum of squares for a given set of parameters \( \rho, \lambda, \alpha \), we can minimize the sum of squared residuals for each pool separately, performing a one dimensional minimization over a single \( \beta \) parameter at a time. We have thus reduced the solution of a problem in 1,159 dimensions to the solution of a series of smaller problems in 4 dimensions. The optimization now becomes feasible.
<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>2.9716</td>
</tr>
<tr>
<td>$\beta$</td>
<td>4.0237</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.6615</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.0336</td>
</tr>
<tr>
<td>$\hat{\sigma}^2$</td>
<td>.00158</td>
</tr>
<tr>
<td>$S_T(\hat{\theta})$</td>
<td>142.608</td>
</tr>
<tr>
<td>$\text{var}(w_{it})$</td>
<td>.00181</td>
</tr>
<tr>
<td>$\text{var}(e_{it})$</td>
<td>.00158</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.126</td>
</tr>
<tr>
<td>$\text{var}(w_t)$</td>
<td>.00025</td>
</tr>
<tr>
<td>$\text{var}(e_t)$</td>
<td>.00002</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.911</td>
</tr>
</tbody>
</table>

Table 1: Parameter estimates from NLLS estimation

4.3 Results

To test the new estimation procedure, and to derive results needed for hypothesis testing later, the basic model with only 4 parameters was estimated using NLLS. The results are shown in table 1. These estimated values and standard errors are similar to those reported in Stanton [38]. As in Stanton [38], it is likely that small sample distributions of both parameter estimates and test statistics are different from their asymptotic counterparts.

In the extended model, with heterogeneous $\beta$ parameters, we are estimating 1,159 parameters. The total number of residuals is $78 \times 1,156 = 90,168$. From equation 12, in order to calculate standard errors for each parameter, we need to invert the $1,159 \times 1,159$ matrix

$$
\left[ \sum_{t=1}^{T} \frac{\partial f_t}{\partial \theta} \frac{\partial f_t}{\partial \theta'} \right].
$$

However, if we are only interested in knowing standard errors for the parameters $\alpha, \rho$ and $\lambda$, and not for each individual $\beta_t$, we can reduce the scale of the calculation by using the partitioned inverse formula. Partition the matrix as

$$
\left[ \sum_{t=1}^{T} \frac{\partial f_t}{\partial \theta} \frac{\partial f_t}{\partial \theta'} \right] = \begin{bmatrix} A & B \\ B' & C \end{bmatrix},
$$

where $A$ is $3 \times 3$, $B$ is $3 \times 1,156$, and $C$ is a $1,156 \times 1,156$ diagonal matrix. The top $3 \times 3$ submatrix of the inverse is

$$
\left[ A - BC^{-1}B' \right]^{-1}.
$$
<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>2.9605</td>
<td>0.0094</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.7792</td>
<td>0.0069</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.0447</td>
<td>0.0003</td>
</tr>
<tr>
<td>$\hat{\sigma}^2$</td>
<td>.000153</td>
<td></td>
</tr>
<tr>
<td>$S_T(\hat{\beta})$</td>
<td>137.807</td>
<td></td>
</tr>
<tr>
<td>$\text{var}(w_{it})$</td>
<td>.00181</td>
<td></td>
</tr>
<tr>
<td>$\text{var}(e_{it})$</td>
<td>.00153</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.156</td>
<td></td>
</tr>
<tr>
<td>$\text{var}(w_t)$</td>
<td>.00025</td>
<td></td>
</tr>
<tr>
<td>$\text{var}(e_t)$</td>
<td>.00002</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.912</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: Parameter estimates for extended model

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1156</td>
<td>3.154</td>
<td>1.760</td>
<td>1.326</td>
<td>0.0001</td>
<td>11.597</td>
</tr>
</tbody>
</table>

Table 3: Individual estimated $\beta$ values

Since $C$ is diagonal, $C^{-1}$ is also diagonal, with each non-zero element being the inverse of the corresponding element in $C$. Thus we do not need to store or calculate the entire inverse covariance matrix just to calculate the top left corner, reducing the storage and computational requirements to a manageable level. Using this simplification, the extended model was estimated from the same data as above. The estimates for parameters $\alpha$, $\rho$ and $\lambda$ are shown in table 2. Summary statistics for the 1,156 estimated $\beta$ values are reported in table 3, and the empirical distribution function for the parameters is shown in figure 3. In both the basic and the extended model the estimated average prepayment cost level is very high. For the basic model with no heterogeneity the parameter estimates reported in table 1 imply a mean prepayment cost of 42% of the remaining principal balance on the loan. For the heterogeneous model, combining the $\alpha$ estimate from table 2 with the average $\beta_t$ value from table 3 gives a mean cost of 48% of the remaining principal balance. Since direct monetary costs associated with refinancing a mortgage are much lower than this, typically not exceeding 7%, these costs seem rather high. There are several possible explanations. The first is that the transaction costs being considered are not merely direct monetary costs, but also include non-monetary costs associated with the inconvenience of actually going to
Figure 3: Empirical distribution of heterogeneous $\beta$ parameter. Histogram shows proportion of pools in sample with estimated $\beta$ falling within $\pm$ 0.25 of value on axis. Line shows empirical c.d.f.
<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>2.5782</td>
</tr>
<tr>
<td>$\beta$</td>
<td>6.5883</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.5749</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.0607</td>
</tr>
</tbody>
</table>

Table 4: Parameter estimates with 10 year investment horizon

the bank, filling out forms etc. However, the level is too high to be easily explained by this alone. Another possible explanation is that the optimal exercise calculations performed in this analysis all assume a 30 year planning horizon for each investor. While each mortgage-holder knows that he or she may prepay for interest rate related or exogenous reasons before that date, each has a non-zero probability of holding the mortgage for 30 years. In reality, people often have much shorter planning horizons. If a mortgage-holder plans to move within 7 years, it will take a lot more to persuade him or her to refinance the mortgage than if there is a possibility of staying for 30 years. Intuitively this is because the fixed cost of refinancing must be amortized over a shorter period. Thus if mortgage-holders have shorter planning horizons than 30 years, they will prepay less often than if their planning horizon is 30 years. Estimating the model assuming a 30 year horizon will result in estimated cost distribution parameters that imply higher costs on average than are actually faced by mortgage-holders, in an attempt to match the data. To see whether this can explain the observed cost levels, the prepayment model was estimated again assuming a 10 year investment horizon. The results are shown in table 4.

The estimate of 0.5749 for $\rho$ is similar to that obtained assuming a 30 year horizon. It implies an average time before prepayment of 1 year 9 months. The estimate of 0.0607 for $\lambda$ says that the probability of prepayment for exogenous reasons in a given year is approximately 5.9%, somewhat higher than when a 30 year horizon was used. The average cost level has dropped from 42% with a 30 year horizon to only 28%. Although this accounts for some of the excessive transaction costs, it is not sufficient by itself. An alternative explanation, believed by many mortgage traders, is that a sizable number of mortgage holders never prepay under any circumstances. These mortgage holders have an effective transaction cost close to 100%, so including them in the overall population increases the average estimated transaction cost. In addition, though the beta distribution has a fairly flexible form, once
we have imposed the restriction that there must be a large weight on the highest cost levels,  
continuity of the density function implies that there will also be a relatively high weight  
associated with slightly lower cost levels. This again results in a higher estimated mean  
transaction cost.

4.4 Statistical Significance of Heterogeneity

We wish to test the hypothesis that there is no heterogeneity among pools. More formally,  
the statistical hypothesis to be tested is  

\[ H_0 : \beta_1 = \beta_2 = \ldots = \beta_{1,156}. \]  

(21)

We can construct a test statistic for this hypothesis as a generalization of the \( F \) or \( \chi^2 \) test  
used in OLS.\(^6\) We can use  

\[ \frac{T - K}{q} \frac{S_T(\bar{\theta}) - S_T(\hat{\theta})}{S_T(\hat{\theta})} \sim F(q, T - K) \]  

(22)

or  

\[ (T - K) \frac{S_T(\bar{\theta}) - S_T(\hat{\theta})}{S_T(\hat{\theta})} \sim \chi^2_q, \]  

(23)

where \( q \) is the number of restrictions, \( T \) is the total sample size and \( K \) is the number of  
parameters being estimated. For the unconstrained model, the sum of squares is 137.807.  
For the constrained model (with a common \( \beta \) value for all pools), the sum of squares is  
142.608. The \( \chi^2 \) test statistic is thus given by  

\[ \chi^2_{1155} = \frac{(1156 \times 78) - 1159}{142.608 - 137.807} \]  

\[ = 3101. \]

This is significant at the 1% level (critical value for \( \chi^2_{1155} \) is 1270), so we can reject the  
hypothesis that there is no heterogeneity among the pools, though we must note that the \( \chi^2 \)  
distribution of the test statistic is a large sample result, and the sample size here is only 78.

5 Residual Analysis

Although the parameter estimates are consistent under fairly general assumptions about the  
covariance structure of the error terms \( u_i \), the formulae used to calculate standard errors

\(^6\)See Amemiya [1], page 136.
and the distributions of the test statistics are correct only if the errors are i.i.d. There are reasons to believe that this may not be the case. First, if the model is missing a significant explanatory factor, this should be reflected in correlation between contemporaneous residuals in different pools. Since the residual for pool 1 month 1 is $u_1$, that for pool 2 month 1 is $u_{79}$ etc., this will show up as a non-zero autocorrelation coefficient of order 78 (and 156, 234 etc.). Second, any such misspecification will also probably lead to systematic patterns in the errors for each individual pool, and hence to serial correlation which should show up as a non-zero autocorrelation coefficient of order 1. Finally, the level of the variance may not be constant over time or across pools. It may vary with such variables as pool size and time since issue. If any such patterns exist, we could account for them by performing the NLLS analogue to feasible GLS estimation. We would form a new set of residuals as sums of linear multiples of the original set, taking into account any heteroscedasticity or cross-correlation to ensure that the new set had no cross-correlation and was homoscedastic.

To check for the presence of cross-sectional or cross-temporal correlation in residuals, the residuals were regressed against lagged residuals. The results are reported in table 5. While some of the coefficients are significantly different from zero, their magnitude is small. While this does not necessarily imply that correcting for such dependence would make no difference to the parameter estimates obtained via NLLS, correcting the model's predictions by adding the fitted values from the regressions would make little difference to predicted prepayment levels. This is supported by the small $R^2$ values obtained. The seven regressions reported in Table 5 show that using lags 1 – 5 plus the contemporaneous residual from the next pool allows us to explain under 0.2% of the variation in the model's prediction errors.

As a further test for serial correlation, rather than calculating a single average correlation coefficient using all residuals, as in table 5, 1,156 first order serial correlation coefficients were calculated for each pool individually. Summary statistics for the results obtained are shown in table 6. The sample mean is not significantly different from zero ($s/\sqrt{n} = 0.00459$), though there seems to be some variation between the values for individual pools. Under an assumption of bivariate normality, if the population correlation coefficient for two series of size $T$ is $\rho$, and the sample value is $r$, Fisher's $z$ transform,

$$z = \frac{\sqrt{T-3}}{2} \log \frac{(1 + r)(1 - \rho)}{(1 - r)(1 + \rho)}$$  \hspace{1cm} (24)
Table 5: Residual Regressions. Coefficients marked with an asterisk are significantly different from 0 at 1%.

<table>
<thead>
<tr>
<th>Regression</th>
<th>$y_{k-1}$</th>
<th>$y_{k-2}$</th>
<th>$y_{k-3}$</th>
<th>$y_{k-4}$</th>
<th>$y_{k-5}$</th>
<th>$y_{k-78}$</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.008</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00007</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-0.007</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.00004</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>-0.02*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0004</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.016*</td>
<td>-</td>
<td>-</td>
<td>0.0003</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.015*</td>
<td>-</td>
<td>0.0002</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.024*</td>
<td>0.0006</td>
</tr>
<tr>
<td>7</td>
<td>0.007</td>
<td>-0.007</td>
<td>-0.02*</td>
<td>-0.016*</td>
<td>-0.015*</td>
<td>0.02*</td>
<td>0.0016</td>
</tr>
</tbody>
</table>

Table 6: Individual estimated first order autocorrelation coefficients has asymptotically a standard normal distribution.\(^7\) Figure 4 shows the distribution of the estimated autocorrelation coefficients, standardized according to equation 24. The standard normal c.d.f. is shown for comparison. However, since we have no reason to assume bivariate normality, we shall not use this as the basis for a statistical test. Instead, write

$$\nu_i = \text{cov}(v_{it}, v_{it-1})$$  \hspace{1cm} (25)

for the first order serial covariance for pool $i$. We wish to test the hypothesis

$$H_0 : \nu_1 = \nu_2 = \ldots = \nu_{1,156}.$$  \hspace{1cm} (26)

We can use the $\chi^2$ test defined in equation 23 above, after noting that the 1,156 estimated autocovariances minimize the sum of squares

$$S(\nu_1, \nu_2, \ldots, \nu_{1,156}) = \sum_{i=1}^{1,156} \sum_{t=2}^{78} (v_{it}v_{it-1} - \nu_i)^2.$$  \hspace{1cm} (27)

The unrestricted sum of squares is 0.6794. The restricted sum of squares (after imposing the restriction that $\nu_1 = \nu_2 = \ldots = 0$) is 0.6895. The number of observations, $T$, is 89,012. The number of parameters being estimated and the number of restrictions are both 1,156. The

\(^7\)See Freund and Walpole [20], page 444.
Figure 4: Empirical distribution of standardized autocorrelation coefficients. Histogram shows proportion of pools in sample with standardized coefficients falling within ± 0.1 of value on axis. Solid line shows empirical c.d.f.; dashed line shows standard normal c.d.f.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>0.0014</td>
<td>0.0003</td>
<td>4.692</td>
<td>0.00004</td>
</tr>
<tr>
<td>INVSIZE</td>
<td>46.97</td>
<td>19.96</td>
<td>2.354</td>
<td>0.0186</td>
</tr>
<tr>
<td>TIME</td>
<td>-1.02e-06</td>
<td>1.98e-06</td>
<td>-0.515</td>
<td>0.606</td>
</tr>
<tr>
<td>SHORT</td>
<td>5.12e-05</td>
<td>2.25e-05</td>
<td>2.274</td>
<td>0.0230</td>
</tr>
<tr>
<td>LONG</td>
<td>-4.30e-05</td>
<td>3.16e-05</td>
<td>-1.359</td>
<td>0.174</td>
</tr>
<tr>
<td>LAGSQUARE</td>
<td>0.0838</td>
<td>0.0034</td>
<td>24.67</td>
<td>0.000001</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>0.008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>2.01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Regression of squared residuals against explanatory variables

$\chi^2_{1,156}$ test statistic calculated from equation 23 is 1.317, which is significant at 1% (though subject to the small sample caveat mentioned earlier). In principle, we could correct for this autocorrelation pool by pool, though we do not do so here.

We now want to test whether the variance of $u_t$ is constant. It might depend on such variables as pool size and time since pool issue, as well as on the expected prepayment level for each month, which will be related to the interest rate level. To test this, $u_t^2$ was regressed against variables INVSIZE, TIME, SHORT, LONG and LAGSQUARE, where

$$
\text{INVSIZE} = 1/(1 + \text{Remaining balance}),
\text{TIME} = \text{Time since issue},
\text{SHORT} = \text{Ibbotson one month T-Bill return},
\text{LONG} = \text{Yield on newly issued 20 yr. Treasury Bonds}.
\text{LAGSQUARE}_t = u_{t-1}^2.
$$

The form of variable INVSIZE derives from the fact that under a simple prepayment model the variance should be proportional to the inverse of the number of mortgages in the pool. The 1 in the denominator is to prevent any division by zero. The form for the variable TIME is to account for seasoning. According to the seasoning hypothesis, mortgages are less likely to prepay in early months than in later months, all else being equal. If there is no prepayment, the variance will also be low. Results are shown in table 7. The only explanatory variables significant at 1% are the constant term and the lagged squared residual. Very little of the variance in the residuals can be explained by movements in the exogenous variables used,
though any changes may be somewhat persistent. Variation in the explanatory variables (including the lagged squared residual) explains only 0.7% of the total variation in the squared residuals.

5.1 Dependence on Other Economic Factors

Many economic factors beyond interest rates are often used as inputs to prepayment prediction models. To see whether the model described in this paper is missing a major predictor of mortgage prepayment rates, and also to test for seasonality in prepayment (which has not been included in the model specification so far), the errors in the fitted prepayment rates from the model were regressed against the variables LONG, the par yield on newly issued 20 year Treasury bonds,\textsuperscript{8} SPREAD, the difference between this rate and the annualized return on one month T-Bills,\textsuperscript{9} START, the monthly number of housing starts,\textsuperscript{10} UNEMP, the unemployment rate in the civilian labor force,\textsuperscript{11} IPGRO, the growth in industrial production,\textsuperscript{12} and SUMMER, an indicator variable taking on the value 1 for months May – August inclusive, 0 otherwise. The regression equation to be estimated is

\[
u = \beta_0 + \beta_1 \text{LONG} + \beta_2 \text{SPREAD} + \beta_3 \text{START} + \beta_4 \text{UNEMP} + \beta_5 \text{IPGRO} + \beta_6 \text{SUMMER} + \epsilon.
\] (28)

OLS regression results are shown in table 8. Three of the regression coefficients (excluding the constant term) are significant at 1%. However, only 0.2% of the variation in the dependent variable is explained by variation in the independent variables. To see whether correcting for the results of this regression by adding the fitted values to the model’s predicted prepayment rates would make much difference, table 9 shows the size of this correction for an increase of 1 standard deviation in each of the explanatory variables (except for the indicator variable SUMMER, where the difference between SUMMER = 0 and SUMMER = 1 is given). The mean prediction error is 0.00046, with a sample standard deviation of 0.039. The mean predicted prepayment level is 0.18 with a sample standard deviation of 0.04. Correcting for

\textsuperscript{8}Source: Salomon Brothers.
\textsuperscript{9}Source: Ibbotson Associates.
\textsuperscript{10}Seasonally adjusted, thousands. Source: Department of Commerce, Census Bureau.
\textsuperscript{11}Seasonally adjusted, percent. Source: Department of Labor, Bureau of Labor Statistics.
\textsuperscript{12}Seasonally adjusted, percent. Source: Board of Governors, Federal Reserve System, statistical release G12.3.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard Error</th>
<th>t-statistic</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTANT</td>
<td>0.0036</td>
<td>0.0013</td>
<td>2.722</td>
<td>0.007</td>
</tr>
<tr>
<td>LONG</td>
<td>8.97e-05</td>
<td>0.0001</td>
<td>0.863</td>
<td>0.388</td>
</tr>
<tr>
<td>SPREAD</td>
<td>0.00087</td>
<td>0.0001</td>
<td>8.132</td>
<td>0.00001</td>
</tr>
<tr>
<td>START</td>
<td>2.203e-06</td>
<td>1.071e-06</td>
<td>2.056</td>
<td>0.0398</td>
</tr>
<tr>
<td>UNEMP</td>
<td>-0.0016</td>
<td>0.00022</td>
<td>-7.274</td>
<td>0.00002</td>
</tr>
<tr>
<td>IPGRO</td>
<td>9.89e-05</td>
<td>0.00024</td>
<td>0.417</td>
<td>0.676</td>
</tr>
<tr>
<td>SUMMER</td>
<td>0.0025</td>
<td>0.00028</td>
<td>8.747</td>
<td>0.00001104</td>
</tr>
</tbody>
</table>

Table 8: Regression of prepayment prediction errors against possible explanatory variables

<table>
<thead>
<tr>
<th>$R^2$</th>
<th>0.0023</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2$</td>
<td>0.039</td>
</tr>
<tr>
<td>Durbin-Watson</td>
<td>1.990</td>
</tr>
</tbody>
</table>

Table 9: Effect on predicted prepayment rates of increasing each explanatory variable by one standard deviation (1 for the indicator variable SUMMER)
<table>
<thead>
<tr>
<th>$\kappa$</th>
<th>$\mu$</th>
<th>$\sigma$</th>
<th>$q$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\rho$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.29</td>
<td>0.079</td>
<td>0.11</td>
<td>-0.12</td>
<td>2.96</td>
<td>4.22</td>
<td>0.68</td>
<td>0.028</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.004)</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>0.60</td>
<td>0.079</td>
<td>0.11</td>
<td>-0.12</td>
<td>2.95</td>
<td>4.23</td>
<td>0.75</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.009)</td>
<td>(0.01)</td>
<td>(0.007)</td>
<td>(0.0002)</td>
</tr>
<tr>
<td>0.29</td>
<td>0.040</td>
<td>0.11</td>
<td>-0.12</td>
<td>0.57</td>
<td>6.28</td>
<td>0.48</td>
<td>0.110</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.007)</td>
<td>(0.04)</td>
<td>(0.002)</td>
<td>(0.0007)</td>
</tr>
<tr>
<td>0.29</td>
<td>0.079</td>
<td>0.20</td>
<td>-0.12</td>
<td>0.81</td>
<td>1.97</td>
<td>0.64</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.007)</td>
<td>(0.008)</td>
<td>(0.003)</td>
<td>(0.0004)</td>
</tr>
</tbody>
</table>

Table 10: Estimated parameter values for different interest rate process parameters

A 1 standard deviation change in the independent variables thus makes little difference to the model's prepayment predictions.

5.2 Interest Rate Process

The results above indicate that using the CIR interest rate process allows us to capture most of the variation in prepayment rates caused by changes in interest rates. It is nevertheless interesting to know how sensitive the prepayment model's parameter estimates are to the parameters of the interest rate model. Therefore the model was estimated using a grid size of 100 for several sets of CIR parameter values. The results are shown in Table 10. The parameter estimates do vary somewhat. Increasing $\kappa$, the speed of adjustment, from 0.3 to 0.6 has little impact. Reducing $\mu$, the long run mean interest rate, from 8% to 4% reduces the estimated mean transaction cost level from 41% to 8% of the remaining principal balance. Increasing $\sigma$, the volatility of interest rates, from 0.11 to 0.2 changes the estimated values of $\alpha$ and $\beta$, but the mean transaction cost level remains at 29% of the remaining principal balance. Changing the interest rate process can therefore have a significant impact on the parameter estimates for a given set of prepayment rates. Conversely, this implies that changing the interest rate process and keeping the parameters constant will produce significant changes in predicted prepayment rates. This is in contrast to purely empirical prepayment models, where a change to the interest rate process does not in general affect predicted prepayment, though it does affect asset and liability values through the valuation model used.

Table 11 shows the result of jointly estimating both the parameters in the prepayment
<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu$</td>
<td>0.0785</td>
<td>0.0003</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>2.899</td>
<td>0.007</td>
</tr>
<tr>
<td>$\beta$</td>
<td>4.135</td>
<td>0.012</td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.651</td>
<td>0.0032</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.030</td>
<td>0.0002</td>
</tr>
<tr>
<td>$\hat{\sigma}^2$</td>
<td>0.00158</td>
<td></td>
</tr>
<tr>
<td>$S_T(\hat{\theta})$</td>
<td>142.606</td>
<td></td>
</tr>
<tr>
<td>$\text{var}(w_{it})$</td>
<td>.00181</td>
<td></td>
</tr>
<tr>
<td>$\text{var}(e_{it})$</td>
<td>.00158</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.126</td>
<td></td>
</tr>
<tr>
<td>$\text{var}(w_t)$</td>
<td>.00025</td>
<td></td>
</tr>
<tr>
<td>$\text{var}(e_t)$</td>
<td>.00002</td>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
<td>.912</td>
<td></td>
</tr>
</tbody>
</table>

Table 11: Parameter estimates from NLLS estimation of CIR/prepayment model

model and parameter $\mu$ in the CIR interest rate model.\textsuperscript{13} The parameter estimates are similar to those obtained with a fixed $\mu$. The reported standard errors are also similar, indicating that treating the interest rate parameters as constant may not significantly bias the reported standard errors.

6 Valuation

Most quoted mortgage-backed security prices are so-called “generic” prices, rather than being pool-specific. A contract to deliver a certain face value of mortgage-backed securities on a particular date allows the seller wide latitude in choosing which specific pools are actually delivered. This gives him the option to choose the least valuable pools for delivery. Generic prices are the prices at which these contracts are written. They take into account the existence of this option. However, the seller of a specific face amount of mortgage-backed securities needs to value each individual pool to determine which is the least valuable, and hence is the optimal pool to deliver to satisfy the terms of the contract. In addition,

\textsuperscript{13}Estimation of all four parameters $\kappa, \mu, \sigma$ and $q$ from prepayment data would be impossible, as only a three dimensional subspace of the four parameters is identified by asset prices. In practice (keeping parameter $q$ fixed to avoid this problem) estimation of more than one of the remaining three interest rate parameters using prepayment data proved impossible. After minimizing the objective function over any one parameter, the other two parameters did not change from their initial values. Because of this, only one interest rate parameter was estimated along with the prepayment parameters.
mortgage-backed security holders require pool-specific rather than generic prices to value their portfolios. It is thus important to be able to determine a price for a mortgage-backed security backed by a specific individual mortgage pool. Wall Street traders recognize that pools are heterogeneous, and therefore have different values. They talk about “fast pay” and “slow pay” pools, but adjustments made to account for this heterogeneity are usually ad hoc corrections rather than arising from a structural model of mortgage pool heterogeneity. One such correction involves scaling the difference between the generic price and par by the ratio of the remaining balance on the pool to the average remaining balance on pools of the same type.

Stanton [38] presented some preliminary mortgage-backed security pricing results. However, the valuation methodology presented there is not adequate for the extended model. It works only if we assume that every pool is identical. It does not use the individual prepayment history of a pool to learn more about the specific characteristics of that pool. We now present a methodology for deriving pool-specific mortgage-backed security prices which is consistent with the rational prepayment model described earlier.

To price a mortgage-backed security when there is unobservable heterogeneity concerning the type of the pool backing the security, expectations of future prepayment should be formed conditional on all information available, including the history of prepayment for the pool. The expectations used in Stanton [38] were formed conditional only on the history of interest rates. No pool-specific prepayment information was used. There are actually two possible sources of between pool transaction cost heterogeneity. Two pools may have the same $\alpha$ and $\beta$, but different initial draws from the beta distribution defined by these parameters. They may alternatively have different values of $\alpha$ and $\beta$. It is numerically infeasible to deal with
the first source of heterogeneity. We shall concentrate on the second source of heterogeneity by making two simplifying assumptions.

**Assumption 1** The initial transaction cost distribution for a pool exactly matches the beta distribution defined by $\alpha$ and $\beta$.

**Assumption 2** Each month the ratio of actual to expected prepayment is the same for each transaction cost level.

The value of a mortgage-backed security is calculated by valuing a single mortgage of each possible transaction cost type, and weighting each value by the expected proportion of that type in the pool, conditional on all available information. Stanton [38] showed how to value a single mortgage by using the Crank-Nicholson finite difference algorithm to solve a partial differential equation for asset and liability values. We therefore need only to determine the expected proportion of each cost level in the mortgage pool over time, conditional on all available information. We shall do this in three steps.

1. Calculate the distribution of costs over time, conditional on a given value of $\beta$.

2. Given the distribution of $\beta$, calculate the unconditional cost distribution.

3. Calculate the distribution of $\beta$ over time, conditional on all available information.

---

14 We need to assign a likelihood to every possible initial distribution of costs among the mortgages in the pool. Suppose there are 1,000 mortgages in a pool, and 30 possible transaction cost levels. Each mortgage can take on any one of the 30 cost levels, so the total number of possible arrangements of costs in the pool is $1000^{30} = 10^{90}$.

We can reduce this number, since the order of the assignment is irrelevant. Only the total number of mortgages of each type is important. The number of ways of assigning one of 30 cost values to each of 1,000 mortgages, ignoring the order of the assignment, is the binomial coefficient

$$\binom{1029}{29} = 1.7 \times 10^{56}.$$ 

Each of these possible arrangements has an associated likelihood which needs to be updated each period as more information arrives. If the processing for each arrangement takes $10^{-5}$ seconds, the time required to update all these likelihood values each period would be over $5 \times 10^{39}$ years.

15 We are assuming that investors are risk-neutral with respect to the distribution over possible pool types. This can be justified by arguing that there are many mortgage pools, so the idiosyncratic risk of a single pool can be diversified away.
6.1 Conditional cost distribution

Given the initial values of $\alpha$ and $\beta_i$ for a pool, and the monthly prepayment rates for that pool, assumptions 1 and 2 allow us to determine unambiguously the proportions of each cost type within the pool each month. For a given pool, write $c_j(t)$ for the proportion of the pool with cost level $X_j, j = 1, 2, \ldots, m$ at time $t$.\(^{16}\) By Assumption 1, the initial proportions, $c_j(0)$ are

$$c_j(0) = c_{ij},$$

(29)

defined in equation 8. Let the critical transaction cost level at month $t$ be $X^*_t$ as described in Stanton [38]. Let $j^*$ be the index of the largest cost level $X_j$ satisfying

$$X_j \leq X^*_t.$$  

(30)

The proportion of the pool at time $t$ with transaction costs less than or equal to $X^*_t$ is

$$P_{it} = \sum_{j=1}^{j^*} c_j(t).$$

(31)

The expected proportion of the pool prepaying in month $t$ is

$$\bar{w}_{it} = P_e(1 - P_{it}^*) + P_r P_{it}^*,$$

(32)

where

$$P_e = 1 - e^{-\lambda/12},$$

(33)

$$P_r = 1 - e^{-(\lambda+\rho)/12},$$

(34)

as defined in Stanton [38]. By Assumption 2, the proportion of the pool with cost level $j$ at the start of month $t + 1$ is

$$c_j(t + 1) = \begin{cases} 
\frac{s_j(t)(1-P_e)}{1-\bar{w}_{it}} & \text{if } j \leq j^* \\
\frac{s_j(t)(1-P_r)}{1-\bar{w}_{it}} & \text{if } j > j^*
\end{cases}$$

(35)

Repeated application of equation 35 for $t = 1, 2, \ldots$ gives the monthly proportion of mortgages in the pool of each transaction cost level conditional on a given initial distribution, defined by the value of $\beta_i$. Write the density of transaction cost level $x$, conditional on $\beta_i$ as $g_{it}(x \mid \beta_i).$.\(^{17}\)

\(^{16}\)The subscript for the pool number is omitted for clarity.

\(^{17}\)This depends on the observed history of interest rates and prepayment rates. Explicit dependence is suppressed for clarity.
6.2 Unconditional cost distribution

A given distribution for the parameter $\beta_i$ induces a distribution over the transaction cost proportions in pool $i$. Let the density of transaction cost level $x$, conditional on $\beta_i$ be $g_{it}(x \mid \beta_i)$ as calculated above. Let the posterior distribution function for $\beta_i$ at time $t$ be $F_{it}(\beta)$, with corresponding p.d.f. $f_{it}(\beta)$. Then the unconditional posterior p.d.f. for transaction cost level $x$ is

$$g_{it}(x) = \int_{\beta=0}^{\infty} f_{it}(\beta)g_{it}(x \mid \beta) \, d\beta. \quad (36)$$

For example, at time 0 we have no pool specific information, so the initial distribution for $\beta_i$, $f_{i0}(\beta)$, is the empirical distribution function for $\beta_i$ estimated above and shown in figure 3. The initial distribution of transaction costs conditional on the value of $\beta_i$ is a beta distribution, with p.d.f.

$$g_{i0}(x \mid \beta) = B(\alpha, \beta)x^{\alpha-1}(1 - x)^{\beta-1}, \quad (37)$$

for $0 < x < 1$. Here $B(\alpha, \beta)$ is the beta function defined by

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}. \quad (38)$$

From equation 36, the initial unconditional distribution of transaction costs is therefore

$$g_{i0}(x) = \int_{\beta=0}^{\infty} f_{i0}(\beta)B(\alpha, \beta)x^{\alpha-1}(1 - x)^{\beta-1} \, d\beta. \quad (39)$$

6.3 Updating

Every period, new prepayment information arrives which needs to be used to derive a new posterior distribution for $\beta_i$. To calculate the new distribution we use Bayes' Theorem. This theorem states that if $h(\theta)$ is the value of the prior distribution of $\theta$ at $\theta$, and we observe the value $x$ of a statistic $x$, then the posterior distribution of $\theta$, $\phi(\theta \mid x)$ is given by the formula

$$\phi(\theta \mid x) = \frac{f(\theta, x)}{g(x)} = \frac{h(\theta)f(x \mid \theta)}{g(x)}, \quad (40)$$

where $f(x \mid \theta)$ is a value of the sampling distribution of $x$ given that $\theta = \theta$, $f(\theta, x)$ is a value of the joint distribution of $\theta$ and $x$, and $g(x)$ is a value of the marginal distribution of $x$. In other words, the relative likelihood that $\theta = \theta_1$ versus $\theta = \theta_2$ is multiplied by the relative probability that we observe $x = x$ conditional on $\theta = \theta_1$ versus $\theta = \theta_2$. 

28
In our case, the prior distribution for $\beta_i$ at date $t$ is $f_{i,t-1}(\beta)$. The posterior distribution is $f_{i,t}(\beta)$. The statistic $x$ whose value we observe is the prepayment level for pool $i$ in period $t$, $w_{it}$. To derive the posterior distribution $f_{i,t}(\beta)$ using equation 40, we need to know the distribution of the prepayment level for the pool, conditional on all possible values of the parameter $\beta_i$. The model would tell us this if we knew the number of mortgages in each pool. However, we do not know the number of mortgages in each pool, only the total principal balance. We could assume some fixed mortgage size, but the posterior distribution would depend on this assumed size. Intuitively, the larger the number of mortgages we assume are in a pool, the less likely we are to observe large deviations from the expected prepayment level each month, the less noise there is in the observed signal, and the more information we get from each monthly prepayment level.

We shall instead, therefore, use an empirically derived distribution function. If the number of mortgages in a pool is large, the proportion prepaying each month, conditional on the level of interest rates and the distribution of transaction costs in the pool, should be approximately normally distributed. We shall assume a normal likelihood for $w_{it}$ conditional on the value $\beta_i$, interest rates. The mean of the distribution, conditional on the level of interest rates, is

$$\bar{w}_{it}(\rho, \lambda, \alpha, \beta_i).$$

We also need to know the variance. We showed above that the squared residuals $(\bar{w}_{it} - w_{it})^2$ have little dependence on the size of the pool,\textsuperscript{18} time since issue or interest rates. We shall therefore assume that the variance of the distribution is a constant $\sigma^2$, equal to the estimated value 137.807/90, 168 = 1.53 $\times$ 10$^{-3}$. So

$$w_{it} \sim N(\bar{w}_{it}, \sigma^2).$$  \hspace{1cm} (41)

The likelihood for observation $w_{it}$ conditional on the value $\beta_i$ is the normal p.d.f.

$$L(w_{it} | \beta_i) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{w_{it} - \bar{w}_{it}(\rho, \lambda, \alpha, \beta_i)}{\sigma} \right)^2},$$  \hspace{1cm} (42)

where $\rho, \lambda, \alpha$ take on the values estimated above. Using equation 40, the posterior distribu-

\textsuperscript{18}This is unexpected. We should expect the variance to be proportional to the inverse of the number of mortgages in the pool, itself proportional to the dollar balance in the pool if we assume a constant mortgage size. This lack of dependence was also documented by Beckett and Morris [8] for mortgage pools underlying FNMA mortgage-backed securities.
tion for \( \beta_i \) is given by

\[
f_{it}(\beta) = \frac{f_{i(t-1)}(\beta) \mathcal{L}(w_{it} \mid \beta)}{\int_{b=0}^{\infty} \mathcal{L}(w_{it} \mid b) \, db}.
\]  

(43)

### 6.4 Calculating Pool-Specific Prices in Practice

We have shown how to use transaction cost information to update the distribution of \( \beta_i \) over time, and how to calculate the distribution of transaction costs in a pool given this distribution for \( \beta_i \). To implement this in practice, since we cannot write down a closed form representation for the functions \( f_{it}(\beta) \), we shall therefore use a discrete distribution for both the transaction costs and the possible values of \( \beta_i \). As in Stanton [38], we shall pick \( m_2 \) values for \( \beta_i \) so that the maximum possible difference between the true (empirical) initial cumulative distribution function and the approximation is minimized. This is achieved by letting each weight initially equal \( 1/m_2 \), with value \( b_k \) given by

\[
b_k = F^{-1} \left( \frac{2k - 1}{2m_2} \right),
\]

(44)

where \( F(b) \) is the empirical unconditional probability that \( \beta_i \leq b \). Write \( P_{kt} \) for the probability that \( \beta_i = b_k \) conditional on information available up to time \( t \). Given the observed prepayment level \( w_{it} \), use equation 42 to calculate \( m_2 \) likelihood values for the prepayment level conditional on the possible values for \( \beta_i \). The discrete posterior distribution for \( \beta_i \) is then given by

\[
P_{kt} = \frac{P_{k(t-1)} \mathcal{L}(w_{it} \mid b_k)}{\sum_{j=1}^{m_2} \mathcal{L}(w_{it} \mid b_j)}.
\]

(45)

We now calculate the expected proportion of each possible transaction cost level by forming the discrete analogue to equation 36. If the proportion of cost level \( X_j \) conditional on \( \beta_i \) is \( c_{jit}(\beta_i) \), then given the discrete distribution for \( \beta_i \) defined by \( P_{kt} \), the unconditional expected proportion, \( c_{jt}^* \), is given by

\[
c_{jt}^* = E_t(\text{Proportion with cost } X_j) = \sum_{k=1}^{m_2} P_{kt} c_{jit}(b_k).
\]

(46)

This gives a simple formula for calculating the expected proportions of each cost level in the pool at time \( t \). If the market value of a single mortgage with transaction cost \( X_j \) is \( V_{jit}^a \), calculated as in Stanton [38], the value of a mortgage-backed security backed by the pool is

\[
V_t = \sum_{j=1}^{m} c_{jt}^* V_{jt}^a.
\]

(47)
Thus we have derived a simple methodology for calculating mortgage-backed security prices over time, which consistently incorporates the information revealed by the observed prepayment history of the pool into the valuation. Future work will use this methodology to compare the model's prices with observed pool-specific mortgage-backed security prices.

7 Extensions

We have estimated the prepayment model using prepayment data, and shown how it can produce pool-specific mortgage-backed security prices. These prices potentially provide a means of testing the model. Pool-specific prices calculated as in section 6 can be compared with market prices by the use of non-linear least squares. This results in another set of parameter estimates for the prepayment model. Testing the statistical hypothesis that the true parameter values are those estimated previously using prepayment rates provides a test of the joint hypothesis that the prepayment model is correct, and that the CIR interest rate process adequately describes movements in interest rates, at least as far as these help to predict prepayment rates.

Section 4 noted that the planning horizon of mortgage-holders may not be the assumed 30 years. This results in over-estimating the true transaction costs faced by mortgage-holders. The model could be extended to incorporate heterogeneous planning horizons by imposing a parametric distribution over possible values between 0 and 30 years. The parameters of this distribution could be estimated along with the parameters $\alpha, \beta, \rho, \lambda$ estimated in this paper.

Stanton [38] described the relationship between the valuation of mortgage-backed securities and that of investment vehicles such as bank certificates of deposit (CDs), which also contain embedded options (in this case the option to cash in the investment and reinvest in a higher yielding alternative). As with the prepayment option in a mortgage, these options are not always exercised according to a rational option exercise model with no transaction costs. The optimal surrender strategy of an investor in a CD can be derived in much the same way as we derived the optimal prepayment strategy of a mortgage holder here. Transaction costs, exogenous reasons for surrender, and discrete times between successive surrender decisions can all be modeled in exactly the same way, but the boundary conditions in the partial differential equation would need to be altered to reflect the different type of option. Given
historical surrender data, the parameters in this model could be estimated using GMM or NLLS, and used to make surrender predictions and value the liability of the issuing institution. Even if surrender data were not available, a first step might be to assume the same values for $\alpha, \beta, \rho$ and $\lambda$ as were estimated for mortgages. The basic methodology described here has many potential applications to the study of “irrationally” exercised option-like instruments.

8 Summary

This paper extends the rational mortgage prepayment model developed in Stanton [38] to capture heterogeneity across mortgage pools. The extended model is estimated using the same dataset, and shown to explain the observed data significantly better than the original model with no pool-level heterogeneity. In regressions of the model’s prediction errors against economic variables, only a small fraction of the prediction errors can be explained by current or lagged values of the level of housing starts, interest rates, unemployment, growth in industrial production or a seasonality dummy variable.

The paper describes a consistent methodology for deriving pool-specific mortgage-backed security prices in the face of this heterogeneity, using the information contained in the prepayment history of the pool to update beliefs about the pool’s characteristics, to calculate fully conditional expectations of future prepayment levels and hence to calculate prices that are consistent with the model and the information available to traders.
References


