The Report Committee for Soo Jeong Hong
Certifies that this is the approved version of the following report:

Modeling Prepayment for Mortgage-Backed Security in Korea

APPROVED BY
SUPERVISING COMMITTEE:

Supervisor:

Thomas W. Sager

Tse-min Lin

Modeling Prepayment for Mortgage-Backed Security in Korea

by

Soo Jeong Hong, B.A.

Report

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Master of Science in Statistics

The University of Texas at Austin

August 2013

Abstract

Modeling Prepayment for Mortgage-Backed Security in Korea

Soo Jeong Hong, M.S.Stat.

The University of Texas at Austin, 2013

Supervisor: Thomas W. Sager

The financial crisis of 2007-2009 exposed the credit risks of mortgage-backed securities (MBS). Nevertheless, prepayment remains one of the most important risks of MBS for MBS issuers and purchasers. When mortgages are paid off early, the flow of interest income that MBS purchasers were expecting is terminated. The risk of prepayment must be managed. MBS sellers or purchasers often obtain insurance to protect against this risk. The purpose of this study is to present a model of prepayment rates that is more accurate and precise than other models currently in use. The new model explains the prepayment rate by using the method of panel-corrected standard errors (PCSE). With this method, refinancing incentive, burnout, seasoning, prepayment penalty, and changes in house prices have a substantial effect on the prepayment rate. This study also shows that a feasible generalized least squares (FGLS), which often results in overconfidence for variables, can erroneously identify certain variables as significant. Such variables include the age of mortgage loans with burnout, seasonality, and unemployment rate. Another major finding of this study is that the development of dummy variables for the early months of mortgage pool can provide a better mechanism than the mortgage loan age variable to estimate borrowers' prepayment characteristics. In addition, this study reveals that, contrary to expectation, the global financial crisis did not lead to a significant increase in the refinancing incentive sensitivity after controlling mortgage loan variables.

Table of Contents

List of Tables

Figure

Chapter 1: Introduction

1.1 Prepayment as a risk

Prior to the global financial crisis, which started with Lehman Brothers' bankruptcy in 2008, participants in the market for mortgage-backed securities (MBS) focused on the risk of prepayment of these securities. During and after the crisis, attention was redirected to delinquency risk. However, prepayment is still one of the most important risks that MBS issuers should manage in order to pay MBS purchasers coupon and principal without loss. Prepayment indicates mortgage borrowers' prepayments of the underlying mortgage loans for the MBS and it is generally expressed by its rate, which is the percentage of a pool's outstanding balance at the start of the month that is prepaid in the given month. The prepayment rate is measured by Single Monthly Mortality (SMM) or Conditional Prepayment Rate (CPR). SMM is the percentage of a pool's outstanding balance at the beginning of the month that was prepaid during the month and CPR is the annualized monthly SMM.

Most people might wonder why prepayment is a risk and not a benefit. Those people must have imagined a situation in which they lend their money to a friend without. If the friend returns the money to them earlier than expected, then that results in a more desirable situation for the lender than if the loan is paid on the regular schedule. That is, in individual trades without interest rate, prepayment is difficult to be considered a risk.

However, when it comes to financial organizations, the prepayment clearly becomes a risk factor because they also have a schedule to pay coupons to investors. To better understand prepayment risk, consider the following simple scenario. Let's assume that a financial organization lends $100,000 at a 6% interest rate for a one year amortization period. The borrower is supposed to repay the monthly interest of $500; $6,000 total plus the principal, but when the borrower pays off the loan in the second month, the lender can receive only $1,000 of interest and $100,000 of principal. Assuming a customer deposits $100,000 at 3% interest at the same time when the organization lends the $100,000. The lender should invest the money received from the borrower in a financial product in order to pay the interest for the deposit but usually the shorter investment period causes the lower interest rate and increases the possibility that the cash from the financial product does not cover the interest for the customer's deposit.

This scenario can be applied to MBS. In the case of MBS, if the reinvestment rate of prepayment from mortgage loans is not enough to pay the coupon for MBS purchasers, the MBS issuer can suffer a loss. This is why MBS issuers have to consider prepayment a financial risk to always be hedged promptly because even if issuers are allowed to exercise their call options to transfer prepayment risks to MBS purchasers or impose prepayment penalty on mortgagors, not all prepayment risk can be transferred. Fundamentally that mainly affects MBS valuation. Specifically, if the underlying assets are fixed-rate mortgage loans, MBS issuers pay more attention to cash flows.

Many researchers have proposed models to predict prepayment of mortgage loans for MBS. The early representative statistical models based on pool-level data for prepayment have been suggested by Asay et al. (1987), Chinloy (1991), and Schorin (1992). They used the spread, the ratio of the spread to the market mortgage rate, the age of mortgage loans, the seasonality, and macroeconomic variables to estimate the prepayment rate. Richard and Roll (1989) developed the non-linear multiplicative model employing the ratio of mortgage coupon rate to refinancing rate, pool burnout factor, age for seasoning effect, and months of year for seasonality effect as the explanatory variables. Office of Federal Housing Enterprise Oversight (OFHEO), Schwartz and Torous (1989), and LaCour-Little et al. (2002) suggested a logit model, a proportional hazard model, and a non-parametric kernel regression examining individual-level mortgage loans and maintained that decisions on prepayment and default could be related to each other. Deng and Liu (2009) focused on how the mortgagors' personal traits affect the prepayment rate, and LaCour-Little et al. (2012) analyzed the effect of parameter instability in the prepayment model caused by financially important event over time.

The purpose of this study is to present a more reliable and precise model that explains the prepayment rate by using panel-corrected standard errors (PCSE) method. This article sets the panel datasets which are comprised of 27 panels of MBS pools issued in 2004-2007 and 95 observations in June 2004-April 2012 to model the monthly prepayment rate, SMM. With the PCSE method, refinancing incentive, burnout, seasoning, prepayment penalty, and changes in house prices have a substantial effect on the prepayment rate. Many studies using time-series cross sectional (TSCS) data erroneously ignore the correlation of regression disturbance over time and between subjects, but such effects can lead to biased statistical inference. Beck and Katz (1995) showed that a feasible generalized least squares (FGLS) method gives overconfident standard errors when it is applied to TSCS data generating too small confidence intervals and increasing the probability of a Type I error. In order to control this complex error structure in TSCS analysis, they proposed a more optimal method that uses OLS and a sandwich type estimator of the covariance matrix of the estimated parameters, which is called PCSE. PCSE is robust to the possibility of non-spherical errors. This study successfully employs PCSE to solve contemporaneous correlation and panel heteroskedasticity and shows that FGLS can regard the age of mortgage loans with burnout, seasonality, and unemployment rate as important factors incorrectly.

Another major finding of this study is that the development of dummy variables for the early months can provide a better mechanism than the mortgage loan age variable for predicting borrower's prepayment characteristics. Without controlling other variables, the age variable has more explanatory power than the month dummies. However, when the burnout variable is added to the model that already has the age variable, age loses its explanatory power and the estimated parameter for age is nearly 0.

This study also reveals that the global financial crisis did not give a significant rise to an increase in the refinancing incentive sensitivity. Without controlling other variables, an increase in the slope of SMM with respect to the refinancing incentive is significant but when it comes to the model with the mortgage loan variables, it becomes unimportant. Moreover, its significance in the model without controlling other variables can be caused by the growth of varied alternative financial products through financial technique developments. Rather than by the global financial crisis.

The plan of this study is as follows: Chapter 1.2 will introduce the literature for mortgage loan prepayment rate models. Chapter 2 will describe the data and methodology to be used for prepayment analysis, including a figure of the prepayment rates and introduction to PCSE. Chapter 3 will illustrate the empirical results and draw conclusion.

1.2 Literature review

Models for the prepayment rate that have mainly been developed are statistical models. Using these statistical models, the ability to predict prepayment becomes weaker in severe macroeconomic financial situations because statistical models rely on empirical data over time. Despite these limitations, statistical models have been preferred to theoretical option models by many researchers because of the computational ease in statistical models.

The early representative statistical models for prepayment have been suggested by Asay et al. (1987), Chinloy (1991), and Schorin (1992). Models developed by them share common ground in that they were based on pool data. Asay et al. used the spread, which was the difference between the mortgage coupon and the market mortgage rate as an explanatory variable to measure refinancing incentive that affected prepayment rate using 30-year mortgage loans. They proposed an arctangent function of the spread as an explanatory variable that expresses CPR. Chinloy (1991) inserted the age of mortgage loan into the existing model because the previous model could not explain burnout effect. However, Chinloy concluded that the age variable was not a useful variable because the effect of age was not significant. Shorin (1992) pools 30-year MBS pools for each year and proposed the ratio of the interest rate spread between the contract rate and the market rate over the contract interest rate, the age of mortgage loan, and dummy variables, which expressed seasonality as independent variables. Also, he analyzed several macroeconomics variables but realized the number of monthly existing house purchases was the only variable of which effect was statistically significant.

Richard and Roll (1989) developed a more sophisticated model that was ordered by Goldman Sachs. This non-linear multiplicative model has refinancing incentive, pool burnout factor, age for seasoning effect, and months of year for seasonality effect as the explanatory variables. Richard and Roll used the ratio of the mortgage coupon to the current refinancing rate as a measure of the mortgagor's refinancing incentive arguing that the spread was a poor approximation to the refinancing incentive. They also found that this ratio is one of the important factors that determined the seasoning process.

Models developed by Office of Federal Housing Enterprise Oversight (OFHEO), Schwartz and Torous (1989), and LaCour-Little et al. (2002) examined individual mortgage loans and maintained that decisions on prepayment and default could be related to each other. OFHEO suggested the logit model assuming the error term was distributed as logistic distribution, and Schwartz and Torous fashioned the proportional hazard model of which baseline hazard funtion was assumed to follow log-logistic distribution. LaCour-Little et al. established the prepayment model using non-parametric kernel regression in order to improve the parametric models.

Deng and Liu (2009) studied determinants of prepayment behavior in China analyzing individual mortgage loans. They employed the competing risks hazard model of prepayment and default and found that the Loan-To-Value (LTV) ratio at origination was negatively related with the prepayment and positively related with the default. Also, they revealed mortgagors in higher-income level or in older age levels tended to prepay more and default less and their gender affected neither prepayment risk nor default risk. LaCour-Little et al. (2012) analyzed the effect of parameter instability in the prepayment model caused by financially important event over time. Using mortgage loans originated in the early 1990s compared to the late 1990s, they concluded that the sensitivity of refinancing decisions based on interest rate changes significantly during the period.

Chapter 2: Data and Methodology

2.1 Data

The data used for this research has been drawn from the HF. HF is a Korean public financial company established in 2004 in accordance with the HF Act. The purpose of its establishment is to improve the welfare of the public and facilitate the development of the national economy by increasing the supply of reliable, long-term housing financing. It is a public finance corporation managing funds under the Public Institution Operation Act. HF has undertaken four major businesses including the supply of long-term fixed mortgages, securitizations, housing finance credit guarantees and reverse mortgage guarantees- (2011 Annual Report, HF).

The data contains the percentage rate of prepayment of MBS that HF issued in 2004-2007 and was collected in January 2004-April 2012. From 2004 to 2007, HF handled 27 issuances of MBS of $10.192 billion. Table 2.1.1 shows the issuance summary of MBS analyzed in this research. The number of observations means the number of months that has passed since the corresponding MBS issuance because HF recorded the prepayment rate of MBS at the end of every month. Table 2.1.1 also shows the properties of underlying assets of each pool including issue amount, weighted average mortgage rate, issue data, and the number of observations. The panels referring to each pool are cross-sectional, and the age of the pools are longitudinal. For each individual MBS, the first observation was made on the last business day of the month that the MBS was issued. Thereafter, each MBS was observed on the last business day of each subsequent month. The observations were made monthly, starting on June 30, 2004, and ending on April 30, 2012. Thus, the numbers of observations for each year are as follows: 2004, 644; 2005, 739; 2006, 351; and 2007, 350. The aggregate total of observation for all years was 2,084.

Table 2.1.1 The issuance summary of MBS in 2004-2007

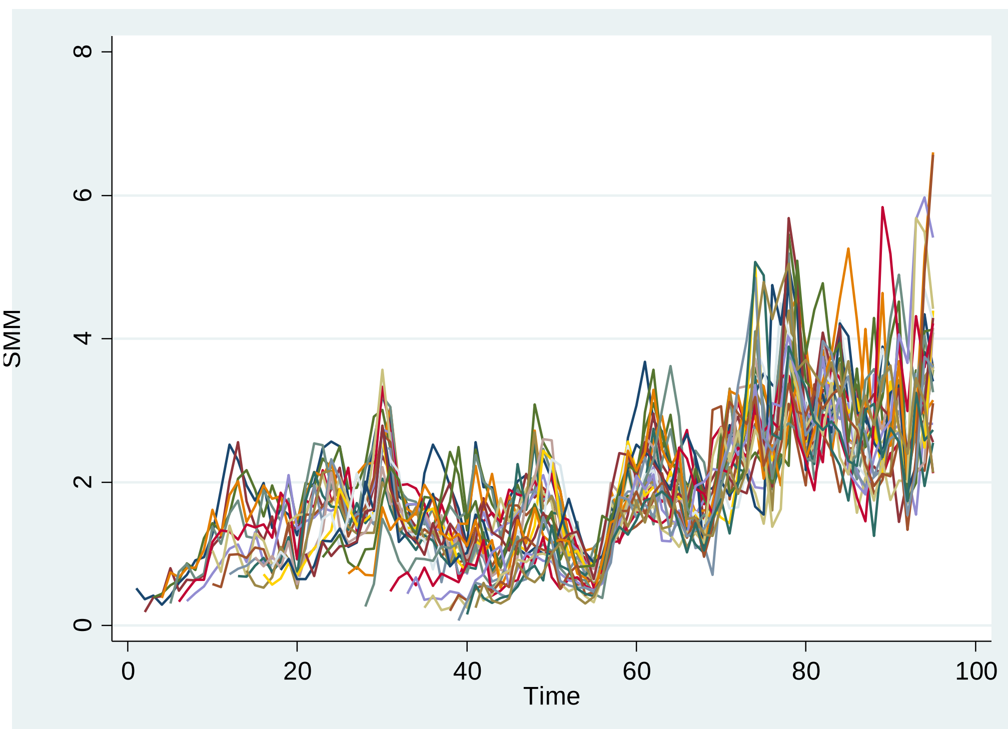
Pool Name	Issue Amount: Billion Korean won	Issue Amount: million US dollars	Weighted Average Mortgage Rate	Issue Date	Number of Observations	Total Number of Observations
MBS2004-01	552.0	520	6.70	6/15/04	95	
MBS2004-02	407.0	383	6.60	7/28/04	94	
MBS2004-03	570.0	537	6.70	8/18/04	93	
MBS2004-04	402.0	379	6.50	9/23/04	92	
MBS2004-05	360.0	339	6.40	10/28/04	91	
MBS2004-06	315.0	297	6.20	11/30/04	90	
MBS2004-07	410.0	386	6.00	12/28/04	89	644
MBS2005-01	408.0	384	5.90	2/24/05	86	
MBS2005-02	405.0	382	5.90	3/23/05	86	
MBS2005-03	467.0	440	5.90	4/28/05	84	
MBS2005-04	453.0	427	5.90	5/25/05	83	
MBS2005-05	459.0	432	6.10	6/27/05	82	
MBS2005-06	454.0	428	6.10	7/27/05	81	
MBS2005-07	484.0	456	6.10	8/30/05	80	
MBS2005-08	400.0	377	6.10	9/29/05	79	
MBS2005-09	331.0	312	6.27	11/10/05	78	739
MBS2006-01	380.0	358	6.34	2/28/06	75	
MBS2006-02	341.5	322	6.58	4/30/06	73	
MBS2006-03	312.5	294	6.47	7/31/06	70	
MBS2006-04	329.6	311	6.42	9/30/06	68	
MBS2006-05	389.5	367	6.24	12/31/06	65	351
MBS2007-01	348.9	329	6.21	02/21/07	63	
MBS2007-02	360.2	339	6.06	04/24/07	61	
MBS2007-03	357.8	337	6.16	07/11/07	58	
MBS2007-04	392.1	369	6.13	08/23/07	57	
MBS2007-05	337.4	318	6.07	09/20/07	56	
MBS2007-06	392.0	369	6.29	10/25/07	55	350
Total	10,818.5	10,192				2,084

Source: HF
Exchange rate: 1US dollar = 1,061.5won (1/2/2013)

Since 2004, the HF has issued Collateralized Mortgage Obligations (CMO) which have various maturity periods ranging from one year to 21 years, indirectly guaranteed by the Korean government. HF's MBS consist of nine tranches based on the same qualities of mortgage loans, which are bullet bonds with different maturities. Except for short-term securities, all the other tranches have a call option to transfer the risk of the mortgage loans to investors. Thus, MBS investors are exposed to the prepayment risk of the underlying assets.

The HF issued MBS seven times in 2004, nine times in 2005, five times in 2006, and six times in 2007. The monetary value of the securitized senior bonds issued in each year was as follows: 3 trillion won ($2.8 billion), 2004; 3.9 trillion won ($3.6 billion), 2005; 1.8 trillion won ($1.6 billion), 2006; and 2.2 trillion won ($2.1 billion), 2007.

Figure 2.1.1 Prepayment rates of underlying mortgage loans for each pool in June 2004–April 2012



The plots in Figure 2.1.1 show the monthly prepayment rates, which is SMM, for each pool by monthly observation. Month 1 represents the first observation month, June 30th 2004, and Month 95 indicates the last observation month, April 30th 2012. The SMM values in each pool basically share similar patterns and the variances have been increasing since around the 68th observation month, January 2010.

2.2 Methodology

The data to be analyzed is the time-series-cross-section (TSCS) data. TSCS data is one of the two types of repeated observation data in which we observe cross-sectional units at more than one time period. If the number of time periods of data (T) is large relative to the number of cross-sectional units (N) the data can be considered TSCS data which can be written as:

$$y_{it} = \boldsymbol{\beta}'\mathbf{x}_{it} + u_{it}; \quad i = 1, \ldots, N; \quad t = 1, \ldots, T \tag{1}$$

where $\mathbf{x}_{it}$ is a vector of one or more exogenous variables and observations are indexed by unit and time. On the other hand, if N is large relative to T, then we can regard the data as panel data. If we regard each MBS pool of the data from HF as a cross-sectional unit, N which is 27 and T which is 95 can characterize TSCS data well.

TSCS data has become a valuable and common resource in social sciences to obtain empirical solutions. Unless we assume the data has a spherical error structure, fitting linear models to TSCS data by ordinary least squares (OLS) will not produce the best linear unbiased estimator (BLUE) because the estimators have incorrect standard errors. TSCS data often contains non-spherical errors because of contemporaneous correlation, panel heteroskedasticity, or serial correlation. TSCS has contemporaneous correlation when errors across cross-sectional units are correlated due to common shocks in a given time period. It can be described as follows:

$$\mathrm{E}(u_{it}, u_{js}) = \begin{cases} \sigma_i^2 \ if \ i = j \ and \ s = t \\ \sigma_{ij} \ if \ i \neq j \ and \ s = t \\ 0 \qquad otherwise \end{cases}$$

Panel heteroskedasticity means the case that the error variance of TSCS data differs across cross-sectional units because of characteristics unique to the units as follows:

$$E(u_{it}, u_{js}) = \begin{cases} \sigma_i^2 \; if \; i = j \; and \; s = t \\ 0 \qquad\qquad otherwise \end{cases}$$

When errors with units are temporally correlated, we can say the data has serial correlation as follows for first-order autoregressive errors:

$$u_{it} = \rho u_{it-1} + \varepsilon_{it}$$

Under the assumption that the variance-covariance matrix (**Ω**) which is used to weight the data is known, the estimator produced by a generalized least squares (GLS) is BLUE with correct standard errors. Since the assumption cannot be satisfied in most cases, Parks (1967) and Kmenta (1986) suggested a feasible generalized least squares (FGLS) estimator by using an estimate $\widehat{\mathbf{\Omega}}$ to improve estimation.

Beck and Katz (1995) showed, however, that the FGLS method gives overconfident standard errors when it is applied to TSCS data generating too small confidence intervals and increasing the probability of a Type I error. FGLS often causes underestimating variability by 50% or more. In order to deal with this complex error structure in TSCS analysis, Beck and Katz (1995) proposed a more optimal method that uses OLS and a sandwich type estimator of the covariance matrix of the estimated parameters, which is called panel-corrected standard errors (PCSE). PCSE is robust to the possibility of non-spherical errors and solves the first two types of non-spherical errors, contemporaneous correlation and panel heteroskedasticity. Serial correlation must be removed before applying the PCSE method.

The OLS method using Equation 1 produces the sampling variability by taking the square root of the diagonal terms of

$$\text{Cov}(\hat{\beta}) = (\mathbf{X'X})^{-1}\{\mathbf{X'\Omega X}\}(\mathbf{X'X})^{-1} \qquad (2)$$

If $\mathbf{\Omega} = \sigma^2\mathbf{I}$, where $\mathbf{I}$ is an $NT \times NT$ identity matrix, the OLS standard errors become the square roots of the diagonal terms of $\hat{\sigma}^2(\mathbf{X'X})^{-1}$ where $\hat{\sigma}^2$ is the usual OLS estimator of the common error variance. To analyze panel models with contemporaneous correlation and panel heteroskedasticity, an $NT \times NT$ block diagonal matrix ($\mathbf{\Omega}$) is used with an $N \times N$ matrix of contemporaneous covariances ($\mathbf{\Sigma}$) along the diagonal. Using the property of consistency of the OLS estimates of Equation 1, we can estimate $\mathbf{\Sigma}$ by using the OLS residuals (e_{it}) as follows:

$$\hat{\mathbf{\Sigma}}_{ij} = \frac{\sum_{t=1}^{T_{ij}} e_{it}e_{jt}}{T_{ij}} \tag{3}$$

where $\hat{\mathbf{\Sigma}}_{ij}$ is an element of $\hat{\mathbf{\Sigma}}$. Thus, we can estimate $\mathbf{\Omega}$ given by $\hat{\mathbf{\Omega}} = \hat{\mathbf{\Sigma}} \otimes I_T$ where $\otimes$ is the Kronecker product and obtain PCSEs which are the square roots of the diagonal terms of

$$\text{PCSE} = (\mathbf{X'X})^{-1}\{\mathbf{X'}\hat{\mathbf{\Omega}}\mathbf{X}\}(\mathbf{X'X})^{-1} \tag{4}$$

The PCSE method can be computed using the R package **pcse** or STATA command **xtpcse**. This study uses the **xtpcse** command in STATA with the panel-specific AR(1) process (**psar(1)**) to handle the serial correlation. The panel-specific AR(1) assumes that a model has N different serial correlations, ρ_i's. In order to control fixed effects for each unit, unit dummy variables are used.

2.3 Variable description

The variables to be examined by PCSE are classified under two groups, factors originated from mortgage loans and factors related to macroeconomic variables. First, four factors from properties of mortgage loans will be studied to find which variables can explain prepayment efficiently, and then, macroeconomic variables will be added to the model that have a high explanatory power to test their effects on prepayment.

The basic full model for this study is as follows:

$$SMM_{it} = \alpha + \beta_1 REF_{it} + \beta_2 REF_{it-1} + \beta_3 REF_{it-2} + \beta_4 REF_{it-3} + \beta_5 B_{it} + \beta_6 AGE_{it} + \sum_{j=1}^{11} \beta_{j+6} MD_{jit} + \sum_{j=1}^{2} \beta_{j+17} MS_{jt} + \beta_{20} Fee_{13it} + \beta_{21} Fee_{37it} + \beta_{22} Fee_{61it} + \sum_{j=1}^{2} \beta_{j+22} HPI_{jt} + \beta_{25} UNEM_t + \beta_{26} UNEMP_t + \beta_{27} INT_{it} + \sum_{j=1}^{26} \beta_{j+27} l_{ji} + u_{it}$$

where SMM_{it} is a single monthly mortality for MBS pool i at time t and $u_{it} = \rho_i u_{it-1} + \varepsilon_{it}$. Chapter 2.3.1 and 2.3.2 will explain each explanatory variable in this equation. The intercept α indicates the y-intercept of the last 27th pool, MBS2007-06 among 27 MBS pools because 26 pool dummy variables, l_{ji}'s, are included to control the fixed pool effect.

2.3.1 Mortgage loan variables

The typical factors of prepayment from mortgage loans are as follows: refinancing incentive, burnout, seasoning, seasonality, and changes in customers' status. Other than these factors, the effect of the prepayment penalty rate is considered as well.

A. Refinancing incentive

Refinancing incentive is oftentimes the most important factor to affect the prepayment of mortgage loans. If the current market rate is lower than the mortgage loan contract rate, borrowers can have the opportunity to refinance the mortgage with a lower interest rate. The borrowers' repayment behavior is affected by the size of the gap between the market rate and the contract interest rate. If the market rate decreases, then the likelihood of the borrowers making a prepayment increases and vice versa. Thus, the expected effect of the refinance incentive is positive.

The refinancing incentive as a consequence of the changes in the cost of refinancing can be measured by the ratio of the mortgage contract rate over the current mortgage loan market rate or the difference between these two rates, the spread. This study employs the first option instead of the spread that many researchers have used. Richard and Roll (1989) contend that the spread is a poor approximation to *A/P* that measures the homeowners' refinancing incentive where *A* is the current mortgage value and *P* is the outstanding principal balance. They use *C/R* where *C* is the mortgage contract rate and *R* is the gross refinancing rate reflecting the cost of refinancing in order to capture the fact that mortgage borrowers show varying delays in responding to refinancing incentives because of differences in processing times by mortgage borrowers. LaCour-Little (2002) also chooses the applied *C/R*, natural log(C/R) to measure the refinancing incentive using a kernel regression. The C/R can be written as:

$$REF_{it} = C_i/R_t$$

where C_i is the weighted average mortgage contract rate of MBS pool i and R_t is market rate at time t. Also, REF_{it-1}, REF_{it-2}, and REF_{it-3} will be examined to consider the time needed to prepare for prepayment after detecting the chance to refinance. The models assume that C_i maintains the initial value during the period of observation.

B. Burnout

Burnout refers to the situation when the current prepayment rate affects the future prepayment rate. Even though the mortgage loan market rate is lower than the contract rate, not all borrowers will initiate a prepayment due to a number of factors, including the possibility of paying transaction costs and issues related to individual credit levels. If the refinancing rate is lower than the market contract rate, then the borrower who refinances first is supposed to incur smaller refinancing costs and have a higher credit level. Thus, if the borrowers who have a higher incentive to refinance are eliminated from the mortgage pool, then in the remaining mortgage pool, the borrowers who have higher refinancing costs and lower credit levels will remain. Therefore, the prepayment rate becomes lower because of heterogeneity across households causing different prepayment behavior; this situation is referred as the "burnout effect." The expected sign of the coefficients for burnout is positive.

Among several measures of burnout (see Hall, 2000), I employ the following formula as a measure of the burnout that LaCour-Little (2002) uses:

$$B_{it} = \sum_{\tau=\tau Orig}^{t} \max\{\log\left(\frac{C_i}{R_\tau}\right), 0\}$$

where C_i is the mortgage contract rate of loan i, R_τ is the prevailing market rate for the mortgage loan at monthly time τ and τOrig is the time of origination. This is the cumulative "moneyness" of REF_{it}.

C. Seasoning

Seasoning is the circumstance that occurs in the first few months following the initiation of the mortgage loan, when prepayments are rare and the prepayment ratio increases over time. The reason for this phenomenon is that the borrowers have a higher financial burden in the early stages of a loan period, and they seek to pursue a stable life. That is, it would be unlikely for borrowers to move again so soon after starting a new mortgage loan.

The Public Securities Association (PSA) model is suggested based on the observation that mortgage prepayments generally increase gradually at the beginning with a constant slope, and stabilize thereafter. Schwartz and Torous (1993) use the age of the mortgage pool. I generate dummy variables for the first gradually increasing part of the PSA curve and add the weighted average monthly age of each pool.

D. Seasonality

Seasonality is the pattern related to moving season, that is, the optimal time during the year to move into a home. Factors that affect seasonality include the start of school, housing construction timing, and weather conditions that are most optimal to move. In general, homeowners move much more frequently in the spring and summer, thus there is a relatively higher occurrence of prepayments during these time periods.

To examine the effect of moving season, I simply test two dummy variables.

$$MS_{1it} = \begin{cases} 1, if\ t = April, May, October, or\ November \\ 0,\ otherwise \end{cases}$$

$$MS_{2it} = \begin{cases} 1, if\ t = March, April, May, September, October,\ or\ November \\ 0,\ otherwise \end{cases}$$

E. Changes in borrowers' status

A number of factors can affect a borrowers' status regarding prepayment behavior such as a change in employment status or divorce. These cases can prompt the selling of a home, whereby the loan is paid off. If the regional economy is booming, prepayments increase; also if borrowers' asset levels are lower, refinancing does not occur frequently. However, this factor is excluded in this study because individual tracking data was not available.

F. Prepayment penalty

Another variable, Fee_{jit} is included to reflect the effect of prepayment penalty rate. Fee_{13it} is 1 only if the weighted average loan age is 13 months, Fee_{37it} is 1 only if the weighted average loan age is 37 months, Fee_{61it} is 1 only if the weighted average loan age is 61 months, and they are 0 in the other cases. These dummies explain the borrowers' prepayment behavior according to the prepayment penalty. The mortgage loan used in this study imposes the prepayment penalty where the ratio decreases from 2% to 1.5% at the individual loan age of 13 months, from 1.5% to 1% at the loan age of 37 months, and from 1% to 0% at the loan age of 61 months (See Table 2.3.1). Because borrowers tend to postpone their prepayment until the prepayment penalty ratio decreases to the next stage, the prepayment ratio is expected to rise at the 13th month, 37th month, and 61st month. Thus, the expected sign of the coefficients for the Fee_{jit} dummies is positive and can be written as:

$$Fee_{13it} = \begin{cases} 1 & \text{, if } Age_{it} \text{ is } 13 \\ 0 & , otherwise \end{cases}$$

$$Fee_{37it} = \begin{cases} 1 & \text{, if } Age_{it} \text{ is } 37 \\ 0 & , otherwise \end{cases}$$

$$Fee_{61it} = \begin{cases} 1 & \text{, if } Age_{it} \text{ is } 61 \\ 0 & , otherwise \end{cases}$$

Table 2.3.1 The prepayment penalty rate of mortgage loans

	Month 0-Month 12	Month 13-Month 36	Month 37-Month 60	Month 61 ~
Prepayment Penalty Rate	1.50%	1.00%	0.05%	0%

2.3.2 Macroeconomic variables

Macroeconomic variables involved in this study are changes in house prices, unemployment rate, and the global financial crisis in order to find the relationship between macroeconomic events and the prepayment that cannot be captured by mortgage loan variables.

A. Changes in house prices

Changes in house prices have a direct effect on the timing and likelihood of a mortgage prepayment as Mattey and Wallace have revealed (2001). Houses are borrowers' assets as well as their collateral in mortgage loans. If a house price increases, in order to realize a profit via the appreciation of the home's value, the borrower can sell the home and pay off the mortgage. Because houses serve as collateral in a mortgage loan, if the price decreases to a certain level, the home loses its value as collateral, so it becomes improbable that the loan will be refinanced. Thus, the expected sign of the estimated coefficient for this variable is positive.

The percentage change of Housing Purchase Price index in nationwide Korea (HPI1) and Seoul (HPI2) collected by the Kookmin Bank are used to examine the effect of changes in house prices on prepayment. HPI measures a relative level of housing purchase price setting the housing purchase price level in June 2011 as 100. I show the effect of these variables without and with controlling the mortgage loan variables.

B. Unemployment rate

The unemployment rate is one of the major macroeconomics indexes that measure the state of the economy. This variable is more known as a determinant of the default rate on mortgage loans but has also been used as a factor that affects prepayment (Campbell and Dietrich, 1983). This study hypothesizes that the relationship of the unemployment rate and prepayment rate is mainly expected to be negative because if losing their job, mortgagors are not supposed to earn a fund surplus easily. The unemployment rate (UNEM) and the percentage change of unemployment rate (UNEMP) are examined and UNEMP is predicted to account for SMM better as sudden changes can be a larger shock and threat to mortgagors to amortize a debt.

C. The global financial crisis

The last macroeconomic variable to be analyzed is the global financial crisis started with Lehman Brothers' bankruptcy triggered by subprime mortgage loans in 2008. The hypothesis linked to this global financial crisis variable is that the sensitivity to the refinancing rate for mortgagors who prepay loans has not changed significantly after the global financial crisis and it is expected to be rejected because this event shook the general economy and the primary losses for customers came from various economic indices and interest rates so that mortgagors became more sensitive to the change of the mortgage rate. This study generates a dummy variable and an interaction term (INT_{it}) between the dummy and the refinance incentive variable, REF_{it-1}. The dummy variable has 0 for the pre-crisis period, January 2004 (1st month, t = 1) – September 2008 (52nd month, t = 52) and has 1 for the post-crisis period, October 2008 (53rd month, t= 53) – April 2012 (95th month, t = 95). If the coefficient for the interaction term is positive, then the result indicates that mortgagors became more sensitive to REF_{it-1} because the coefficient means the difference in slope between two periods.

Chapter 3: Result and Conclusion

3.1 Result

3.1.1 Mortgage loan variables

A. Refinance incentive

The intuition of using the refinancing incentive is that the current market mortgage rate which is lower than the mortgage loan contract rate compels borrowers to refinance the mortgage. The borrowers' repayment behavior is affected by the size of the gap between the market rate and the contract interest rate. If the market rate decreases, then the likelihood of the borrowers making a prepayment increases and vice versa. In order to consider the time to prepare for refinancing after capturing a chance to refinance, I examine the effect of REF_{it}, REF_{it-1}, REF_{it-2}, and REF_{it-3} on prepayment. The models assume that C_i maintains the initial value during the period of observation.

Table 3.1.1.depicts the estimates of the effect of refinancing incentive as a result of changes in the refinance rate. Considering the R^2 values, Model 2 including REF_{it-1} accounts for the prepayment rate, SMM, best. In other words, mortgagors seem to implement their option to prepay one month after making a decision to prepay based on the difference between the cost of prepayment and the cost of maintaining the current mortgage loan. Thus, if REF_{it-1} rises by one, it increases the prepayment rate by 4.39%p at time t without controlling other variables.

Table 3.1.1 The estimated coefficient of refinancing incentive

Model Number	REF_{it}	REF_{it-1}	REF_{it-2}	REF_{it-3}	α	R^2
1	4.44888***	-	-	-	-3.068472***	0.2247
2	-	4.39486***	-	-	-2.943361***	0.2578
3	-	-	4.048373***	-	-2.525853***	0.2424
4	-	-	-	3.800381***	-2.2052***	0.2240

* significant under α = 10%
** significant under α = 5%
*** significant under α = 1%

B. Burnout

The basic concept of burnout comes from the fact that not all borrowers show the same prepayment behavior under the same *REF* because they face different refinancing costs. Table 3.1.2 provides the effect of burnout, B_{it}, on prepayment and indicates that the burnout effect is significant without considering other variables. An increase of 1 in B_{it} causes prepayment to rise 0.25% and the explanatory power is higher than Model 9 that holds $REF_{\text{it-1}}$ and $REF_{\text{it-3}}$ based on the R^2 value.

Table 3.1.2 The estimated coefficient of burnout

Model Number	B_{it}	α	R^2
5	0.2537056***	1.205854***	0.3095

Adding the variable of burnout to Model 1 – Model 4, four other models, Model 6 – Model 9 can be created. Among the models containing the refinancing incentive and the burnout variable, the model that has B_{it} and $REF_{\text{it-1}}$ achieves the highest R^2 value which is 0.4350. Thus, if one wants to use burnout and refinancing incentive only, the choice of B_{it} and $REF_{\text{it-1}}$ seems to be the most efficient.

C. Seasoning

Seasoning effect reflects the phenomenon that prepayment rates tend to be different before and after the MBS pool becomes mature. In order to exhibit the seasoning effect, the PSA model and the age variable are analyzed.

The PSA model suggests two parts of prepayment rate, the constantly increasing part and the flat part. The annualized prepayment rate, which is CPR, rises for 30 months and after it reaches 6%, becomes flat in the PSA model as follows:

$$CPR_t = \begin{cases} \frac{6}{30}t & If\ t \leq 30 \\ 6 & If\ t > 30 \end{cases}$$

where $CPR_t = 100 \times (1 - (1 - SMM_t)^{12})$ and t is the monthly age of underlying loans in a MBS pool. The data for this study shows that the first constantly growing part lasts for 11 months as a result of testing age dummy variables. These dummies can be written as:

$$MD_{jit} = \begin{cases} 1 & \text{, if } Age_{it} \text{ is j} \\ 0 & , otherwise \end{cases}$$

where Age_{it} is the weighted average loan age in MBS pool i. Since the all MBS pools were issued at $Age_{it} = 3$, MD_{1it} and MD_{2it} are dropped in this study. Table 3.1.3 shows the result of testing nine MD_{jit} variables.

Table 3.1.3 The estimated coefficients of month dummy variables

Model Number	Variable	Estimate
10	MD_{3it}	-1.400138***
	MD_{4it}	-1.246042***
	MD_{5it}	-1.15517***
	MD_{6it}	-1.163221***
	MD_{7it}	-1.046103***
	MD_{8it}	-0.9334742***
	MD_{9it}	-0.7590685***
	MD_{10it}	-0.6448787***
	MD_{11it}	-0.374032***
	α	2.153782***
	R^2	.0729

As the PSA curve displays, the estimates corresponding to MD_{jit} variables are increasing by roughly uniform increments over time.

Schwartz and Torous (1993) simply use the age of the mortgage pool. Table 3.1.4 exhibits the consequence of testing the weighted average monthly age of each pool. The effect of the weighted average loan age is significant and one additional monthly age increases the prepayment rate by 0.028%p.

Table 3.1.4 The estimated coefficient of monthly age of MBS pool

Model Number	Age_{it}	α	R^2
11	.0284267 ***	1.06995 **	0.2394

When Model 10 and Model 11 are merged into Model 12, only the effect of Age_{it} is significant and the month dummy variables become insignificant. The R^2 value is 0.2445 for Model 12. However, keeping month dummy variables in the model for further analysis does not seem to be harmful because only they can reveal the initial part of the prepayment rate that shows an increase by small increments until the underlying assets become mature.

Table 3.1.5 The estimates for refinancing incentive, burnout, and seasoning variables

Variable	13	14	15	16	17	18	19	20	21
REF_{it-1}	3.13***	3.42***	-	3.13***	4.36***	-	2.70***	3.00***	-
B_{it}	.10***	-	.21***	.11***	-	.23***	.08**	-	.20***
Age_{it}	.00	.01***	.00	-	-	-	.01**	.00***	.01
MD_{3it}	-	-	-.56***	-	-	-.61***	-	-	-
MD_{4it}	-.92***	-.87***	-.45**	-.96***	-1.39***	-.49***	-	-	-
MD_{5it}	-.85***	-.80***	-.40**	-.89***	-1.31***	-.44**	-	-	-
MD_{6it}	-.88***	-.83***	-.44**	-.92***	-1.33***	-.48***	-	-	-
MD_{7it}	-.77***	-.73***	-.39**	-.81***	-1.19***	-.42**	-	-	-
MD_{8it}	-.69***	-.64***	-.34**	-.72***	-1.08***	-.37**	-	-	-
MD_{9it}	-.55***	-.51***	-.26*	-.59***	-.90***	-.29*	-	-	-
MD_{10it}	-.30***	-.47***	-.26*	-.54***	-.80***	-.29**	-	-	-
MD_{11it}	-.65***	-.27***	-.14	-.32***	-.49***	-.16	-	-	-
α	-1.72***	-.98**	1.34***	-1.69***	-2.70***	1.38***	-1.58***	-1.94***	
R^2	0.5246	0.5126	0.3199	0.5229	0.4404	0.3184	0.4565	0.4468	0.3143

Table 3.1.5 shows that the combination of refinancing incentive, burnout, and seasoning variables. MD_{3it}, is omitted in some models because of multicollinearity caused by adding REF_{it-1} to the model. Since the estimate of Age_{it} is nearly zero or insignificant, further studies will rule out Age_{it} and contain the other three significant variables REF_{it-1}, B_{it}, and MD_{jit}'s. This is because Age_{it} and B_{it}, capture similar traits of SMM even though they are originated from different concepts.

D. Seasonality

The seasonal patterns of prepayment are known to be important because the events of moving that usually involve selling a house can increase SMM. However, two tests show that the moving season dummy variables MS_{1it} and MS_{2it} do not affect SMM significantly. Thus, seasonality is dropped for the further studies.

Table 3.1.6 The estimated coefficient of seasonality

Model Number	MS_{1it}	MS_{2it}	α	R^2
22	.117982	-	1.791068^{***}	0.0305
23	-	.1096798	1.785024^{***}	0.0290

E. Prepayment penalty

The reason why the prepayment penalty effect is considered is that borrowers are expected to postpone their prepayment until the prepayment penalty rate decreases to the next lower level.

As a result of test three dummy variables of Fee_{jit}, the effects of two variables for the 37th month and 61st month are statistically significant as Table 3.1.7 depicts. The mortgage age of 37th month and 61st month affect the refinancing rate significantly positively as they were expected. That is, borrowers tend to postpone their prepayment until the penalty rate goes down to 1%p or 0.5%p. The consequence for Fee_{13it} can be explained by the possibility that the portion of a borrower group who care more about prepayment penalty rate grows over time after the other borrowers who do not care about the prepayment penalty rate as much as the group or by the possibility that borrowers have a better understanding of their mortgage loans over time. Fee_{13it} is ignored in the further study.

Table 3.1.7 The estimated coefficient of prepayment penalty

Model Number	Fee_{13it}	Fee_{37it}	Fee_{61it}	α	R^2
24	.0424969	.1389802**	.4532083***	1.854221***	0.0434

Combining previous models and two significant prepayment penalty variables together, Model 25 can be obtained. Model 25 is the model which does not include insignificant variables and achieves the highest R^2 value. Table 3.1.8 displays Model 25, and macroeconomic variables are added to this model in Chapter 3.1.2.

Starting in 2012, the prepayment penalty rate changed from a decreasing incremental pattern to a sliding method for three years. At this time, the penalties related to making prepayment changed as follows; the period that a penalty is incurred was shortened from five years to three years, and because the sliding method was implemented, the penalty rate was decreased on a daily basis over a three-year period. Thus, people would not pay attention to the prepayment penalty rate from that time point but most remaining loans which were already securitized follow the previous decreasing incremental pattern of prepayment rate, the three *Fee* dummy variables should be important to predict the prepayment for seven years more, which is the average maturity of MBS.

Table 3.1.8 The estimated coefficient of prepayment penalty with other variables

Model Number	Variable	Estimate
25	REF_{it-1}	3.119169***
	B_{it}	.1147218***
	MD_{4it}	-.9406151***
	MD_{5it}	-.8751068***
	MD_{6it}	-.9007971***
	MD_{7it}	-.7988875***
	MD_{8it}	-.7100586***
	MD_{9it}	-.5761317***
	MD_{10it}	-.5229704***
	MD_{11it}	-.3060079***
	Fee_{37it}	.2528372***
	Fee_{61it}	.5436164***
	α	-1.684579***
	R^2	0.5612

3.1.2 Macroeconomic variables

Using Model 25, which explains SMM relatively well, the study further examines the relationship between macroeconomic events and the prepayment that cannot be captured by mortgage loan variables. Figure 2.1.1 shows the possibility that macroeconomic variables improve the prepayment rate model since the plots share the trend to a certain degree over time. Also the R^2 value for Model 25 is 0.5612 which means that we can uncover some missing variables.

A. Changes in house prices

The percentage change of Housing Purchase Price index in Korea nationwide (HPI_1) and Seoul (HPI_2) are used to examine the effect of changes in house prices on prepayment. In Table 3.1.8, Model 26 and Model 27 show the effect of HPI's respectively, and Model 28 and Model 29 are comprised of the variables in Model 25 and the HPI's. Models that contain both HPI_1 and HPI_2 are excluded.

HPI_1 and HPI_2 are significant in all models in Table 3.1.8, but the nationwide index HPI_1 affects the prepayment rate more and explains it better than HPI_2 for Seoul. One percent change in HPI_1 causes an increase in SMM by 0.39% and one percent change in HPI_2 increases SMM by 0.16% without controlling other variables. Under controlling $REF_{\text{it-1}}$, B_{it}, MD_{4it}'s, and Fee_{jit}'s, one percent change in HPI_2 still increases SMM by 0.34% and one percent change in HPI_2 increases SMM by 0.18%. The effects for two variables maintain the similar magnitude regardless of the other mortgage loan variables. Thus, when the housing market can be predicted more precisely, forecasting the prepayment rate becomes easier. If the proportion of mortgagors who live in Seoul is known, it is possible for the mixture of HPI_1 or HPI_2 to improve models efficiently.

Table 3.1.9 The estimated coefficient of changes in house prices

Variable	26	27	28	29
HPI_{1it}	.396046***	-	.3169739***	-
HPI_{2it}	-	.1687088**	-	.1658836***
$REF_{\text{it-1}}$	-	-	2.864258***	2.998941***
B_{it}	-	-	.1184093***	.1305405***
MD_{4it}	-	-	-.8913998***	-.8938903***
MD_{5it}	-	-	-.84323***	-.8428727***
MD_{6it}	-	-	-.867453***	-.8723805***
MD_{7it}	-	-	-.7838137***	-.7875908***
MD_{8it}	-	-	-.6985498***	-.6999557***
MD_{9it}	-	-	-.597584***	-.5960263***
MD_{10it}	-	-	-.5765451***	-.5646042***
MD_{11it}	-	-	-.351937***	-.342158***
Fee_{37it}	-	-	.2635535***	.2572411***
Fee_{61it}	-	-	.5910189***	.5733451***
α	1.764299***	1.837786***	-1.486691***	-1.615314***
R^2	0.0660	0.0372	0.6116	0.5929

B. Unemployment rate

This study set the hypothesis that the unemployment rate and prepayment rate are negatively related because losing a job generates mortgagors that are not supposed to earn fund surpluses easily, and thus examines the unemployment rate (UNEM) and the percentage change of unemployment rate (UNEMP). Although the estimates of the effect for these two variables are negative values as expected, both Model 30 and Model 31 do not sufficiently explain borrowers' prepayment behavior.

Table 3.1.10 The estimated coefficient of unemployment rate

Model Number	$UNEM_{it}$	$UNEMP_{it}$	α	R^2	Prob> χ^2
30	-.1864053	-	2.491108 ***	0.0265	.9507
31	-	-.0091812**	1.846114***	0.0460	.7850

C. The global financial crisis

Lastly this article attempts to reveal a change in sensitivity to the refinancing incentive while mortgagors went through the global financial crisis. The hypothesis which is predicted to be rejected is that the sensitivity to the refinancing rate for mortgagors who prepay loans has not changed significantly after the global financial crisis. In order to test this hypothesis, the interaction term (INT_{it}) between REF_{it-1}.and the dummy variable of which value is 0 for the pre-crisis period (June 2004-September 2008), and of which value is 1 for the post-crisis period (October 2008-April 2012) is set.

Table 3.1.10 shows that the prepayment rate slope with respect to REF_{it-1} changed in the expected direction after the financial crisis. Without controlling other variables, the slope of SMM with respect to REF_{it-1} is 3.22 before the financial crisis occurred and the slope rises to 3.69 after the event in Model 32. Model 33 also illustrates that the slope increases after the financial crisis controlling the significant macroeconomic variable, HPI. This indicates the ratio of the mortgage contract rate to the market mortgage rate can provide more powerful refinancing incentive to mortgage borrowers. However, when it comes to Model 34, INT_{it} turns out to be an unimportant factor with the mortgage loan variables in Model 25. This means that the mortgage loan variables already explain the trait of SMM that is accounted for by INT_{it}.

Table 3.1.11 The estimated coefficient regarding the global financial crisis

Variable	32	33	34
INT_{it}	.4731025***	.6626296***	-.1194024
REF_{it-1}	3.222426***	2.386252***	3.307578***
HPI_{1it}	-	.4435916***	-
B_{it}	-	-	.1252704***
MD_{4it}	-	-	-.9857278***
MD_{5it}	-	-	-.921703***
MD_{6it}	-	-	-.9483834***
MD_{7it}	-	-	-.8458774***
MD_{8it}	-	-	-.7574514***
MD_{9it}	-	-	-.6213556***
MD_{10it}	-	-	-.5645343***
MD_{11it}	-	-	-.3370516***
Fee_{37it}	-	-	.2501769***
Fee_{61it}	-	-	.5562958***
α	-2.057605***	-1.396146**	-1.807123***
R^2	0.3171	0.4167	0.5640

In order to check if the short-term shock that changes the sensitivity to REF_{it-1} exists, five interaction terms between REF_{it-1} and five dummy variables are set up. The first dummy variable value is 1 for the first 3 months (October 2008-December 2008) after the financial crisis, and is 0 for the other period. In the similar manner, the second dummy variable value is 1 for the first 6 months (October 2008-March 2009), and 0 for the other period. The third dummy variable value is 1 for the first 12 months (October 2008-September 2009), the forth one is 1 for the first 18 months (October 2008-March 2010), and the fifth one is 1 for the first 24 months (October 2008-September 2010) and 0 for the other period. The five models including each combination of each interaction tem and REF_{it-1} produce insignificant interaction terms. However, the interaction term for longer period has a smaller p-value. This result implies that the significant increase in the slope of SMM with respect to REF_{it-1} in Model 32 or Model 33 does not come from the global financial crisis but might have come from the growth of varied alternative financial products through financial technique developments.

3.1.3 Comparison to FGLS result

The first five models in Table 3.1.11 display have a variable which can be erroneously regarded as a more important variable using FGLS method allowing AR(1) structure. The results using FGLS show that the estimated standard errors are significantly too small, underestimating variability by more than 50% although the estimators are unbiased. On the contrary, one last result is different from the expectation. Model in which prepayment penalty variables are involved shows that the estimated standard errors by FGLS method are slightly smaller than ones produced by PCSE. The reason for this can be discussed for the future studies.

Table 3.1.12 The results from PCSE and FGLS method

Model Number			B_{it}	Age_{it}	MS_{1t}	MS_{2t}	$UNEM_t$	$UNEMP_t$	Fee_{13it}	Fee_{37it}	Fee_{61it}
21	PCSE	Estimate	.20***	.01	-	-	-	-	-	-	-
		Std.Err.	.054	.006	-	-	-	-	-	-	-
	FGLS	Estimate	.20***	.01***	-	-	-	-	-	-	-
		Std.Err.	.021	.003	-	-	-	-	-	-	-
22	PCSE	Estimate	-	-	.12	-	-	-	-	-	-
		Std.Err.	-	-	.073	-	-	-	-	-	-
	FGLS	Estimate	-	-	.12***	-	-	-	-	-	-
		Std.Err.	-	-	.023	-	-	-	-	-	-
23	PCSE	Estimate	-	-	-	.11	-	-	-	-	-
		Std.Err.	-	-	-	.072	-	-	-	-	-
	FGLS	Estimate	-	-	-	.11***	-	-	-	-	-
		Std.Err.	-	-	-	.023	-	-	-	-	-
30	PCSE	Estimate	-	-	-	-	-.19	-	-	-	-
		Std.Err.	-	-	-	-	.130	-	-	-	-
	FGLS	Estimate	-	-	-	-	-.19***	-	-	-	-
		Std.Err.	-	-	-	-	.041	-	-	-	-
31	PCSE	Estimate	-	-	-	-	-	-01**	-	-	-
		Std.Err.	-	-	-	-	-	.004	-	-	-
	FGLS	Estimate	-	-	-	-	-	-01***	-	-	-
		Std.Err.	-	-	-	-	-	.001	-	-	-
24	PCSE	Estimate	-	-	-	-	-	-	.04	.14*	.45***
		Std.Err.	-	-	-	-	-	-	.082	.082	.092
	FGLS	Estimate	-	-	-	-	-	-	.04	.14*	.45***
		Std.Err.	-	-	-	-	-	-	.084	.084	.092

3.2 Conclusion

This study has presented PCSE as a technique for estimating prepayments using MBS pool-level data. Controlling contemporaneous correlation, panel heteroskedasticity, the PCSE method facilitates more precise estimation for the prepayment rate with five kinds of variables, the refinancing incentive, the burnout, the dummy variables for the initial 11 months as a seasoning variable, changes in the prepayment penalty rate, and changes in housing prices (Model 28). This result suggests the possibility that other studies for modeling prepayment rate using TSCS data produced incorrect results. Utilizing FGLS method, they might have regarded statistically insignificant variables as important factors for the prepayment rate because FGLS method leads to overconfident standard errors. The empirical results in this study show that many insignificant estimated parameters obtained by the PCSE method turn out to be statistically significant in FGLS analysis because of overconfident standard errors. However, only in case of the estimated coefficients for Fee_{jit} dummy variables, their standard errors are slightly smaller using FGLS than ones using PCSE. Analyzing the reasons that cause smaller standard errors by FGLS using TSCS data can be a meaningful for the further research.

Another major finding of this study is that the development of dummy variables for the early months can provide a better mechanism than the mortgage loan age variable for predicting borrower’s prepayment characteristics. Without controlling other variables, the age variable has more explanatory power than the month dummies. However, when the burnout variable is added to the model that already has the age variable, age loses its explanatory power and the estimated parameters for the age is nearly 0.

Lastly, this study reveals that the global financial crisis did not give a rise to an increase in the refinancing incentive sensitivity. Without controlling other variables, an increase in the slope of SMM with respect to the refinancing incentive is significant but when it comes to the model with the mortgage loan variables, it becomes unimportant. Moreover, its significance in the model without controlling other variables can be caused by the growth of varied alternative financial products through financial technique developments not by the global financial crisis.

One limitation of the study is that this study uses pool-level data. Thus, many other methodologies that are hampered by pool data are not applied. If the individual prepayment data is accessible, more varied and precise models can be examined to account for the prepayment rate. In the literature, researchers utilized borrowers' income level, LTV ratios, gender, the length of time until their prepayments occur and so on that are not included in this study. Based on the available data, the respective traits for each MBS pool cannot be verified because large variations in the composition of more than 4,000 loans among the pools can generate factors to determine the individual pool effects.

Also, the model assumes that the weighted average mortgage rate of each pool is fixed during the amortization period but it is actually changing as borrowers make a repayment. Since the people who maintain a high credit level to refinance or pay the higher interest rate compared to the current market rate are expected to make prepayments first, the weighted average rate would decrease gradually. Thus, for a more accurate prediction, the changes of weighted average mortgage rates can be considered.

The other limitation for all statistical models for the prepayment rate is that the models can be more essential when the precise forecast for the future financial market and macroeconomic variables comes along. Thus, this study can be more meaningful if the market mortgage rate, changes in house prices, and changes in mortgagors' behavior according to the information technology development become predictable more efficiently.

The study is that data for only one financial product is involved whereas multiple products are available to borrowers. Following the period for which the data for the present study was collected, new products have been introduced to the market. One of these is an Internet-based mortgage loan. Customers of this product are expected to be more sensitive to the refinancing rate which mainly affects the refinancing incentive. Customers from another product such as a mortgage loan that has both the fixed-rate period and floating-rate period are expected to show the different fashion in prepayment rate. Therefore, more diverse prepayment models need to be established because financial product structures have had the tendency to vary according to the financial technique development.

Appendix

STATA commands

- PCSE analysis

  tsset pool_label time

  xtpcse {a dependent variable} {independent variables}, corr(psar1)

- FGLS analysis

  xtgls {a dependent variable} {independent variables}, corr(psar1)

- Generating a dummy variable

  gen {new variable} = 0

  replace {new variable} = 1 if {condition}

- Generating lagged variables

  gen lag1 = x[_n-1]

  gen lag2 = x[_n-2]

  gen lag3 = x[_n-3]

- Generating a two-way plot

```
twoway (line smm time if pool_name ==200401)(line smm time if pool_name ==200402)(line smm time if pool_name ==200403)(line smm time if pool_name ==200404)(line smm time if pool_name ==200405)(line smm time if pool_name ==200406)(line smm time if pool_name ==200407)(line smm time if pool_name ==200501)(line smm time if pool_name ==200502)(line smm time if pool_name ==200503)(line smm time if pool_name ==200504)(line smm time if pool_name ==200505)(line smm time if pool_name ==200506)(line smm time if pool_name ==200507)(line smm time if pool_name ==200508)(line smm time if pool_name ==200509)(line smm time if pool_name ==200601)(line smm time if pool_name ==200602)(line smm time if pool_name ==200603)(line smm time if pool_name ==200604)(line smm time if pool_name ==200605)(line smm time if pool_name ==200701)(line smm time if pool_name ==200702)(line smm time if pool_name ==200703)(line smm time if pool_name ==200704)(line smm time if pool_name ==200705)(line smm time if pool_name ==200706)
```

Bibliography

Asay, M., Guillaume, F., & Mattu, R. (1987). Duration and convexity of mortgage-backed securities: some hedging implications from a simple prepayment-linked present-value model. In: F. Fabozzi (Ed.), Mortgage-Backed Securities (pp. 103±125). Chicago: Probus.

Beck, N. & Katz, J. (1995). "What to do (and not to do) with Time-Series Cross-Section Data." *The American Political Science Review*, 89(3), 634-647

Bailey, D. & Katz, J. (2011). "Implementing Panel-Corrected Standard Errors in R: The pcse Package." *Journal of Statistical Software*, 42, 1-11.

Campbell, T., & Dietrich, J. (1983). "The Determinants of Default on Insured Conventional Residential Mortgage Loans." *Journal of Finance*, 38(5), 1569-1581

Deng, Y. & Liu, P. (2009). "Mortgage Prepayment and Default Behavior with Embedded Forward Contract Risks in China's Housing Market." *Journal of Real Estate Finance and Economics*, 38(3), 214-240

Hall, A. (2000). "Controlling for Burnout in Estimating Mortgage Prepayment Models." *Journal of Housing Economics*, 9:4, 215-232

Hoechle, D. (2007). "Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence.", *Stata Jounal*, 1-31

Hu, J. (2011). *Appendix A in Asset Securitization: Theory and Prcatice*, John Wiley & Sons (Asia) Pte. Ltd.

Kmenta J. (1986). *Elements of Econometrics*. 2nd edition. Macmillan, New York.

LaCour-Little, M., Marschoun, M., and Maxam, C. (2002). "Improving Parametric Mortgage Prepayment Models with Non-parametric Kernel Regression." *Journal of Real Estate Research*, Vol. 24, No. 3, 2002, 299-326

Mattey, J. & Wallace, N. (2001). "Housing-Price Cycles and Prepayment Rates of U.S. Mortgage Pools." *Journal of Real Estate Finance and Economics*, 23(2), 161-284

Park, Y. and Bang, D. (2011). "An Analysis of the Determinants of Fixed Rate Mortgage Laons' Conditional Prepayment Rates: Using Pool Level Time-Series Regression." *Housing Studies Review*, 19(3), 5-27

Parks, R.. (1967). "Effcient Estimation of a System of Regression Equations When Disturbances Are Both Serially and Contemporaneously Correlated." *Journal of the American Statistical Association*, 62(318), 500-509.

Pennington-Cross, A., (2002). “Patterns of Default and Prepayment for Prime and Nonprime Mortgage.” Office of Federal Housing Enterprise & Oversight Working paper
Richard, S. & Roll, R. (1989). “Prepayment on Fixed-Rate Mortgage-Backed Securities.” *Journal of Portfolio Management*, Spring, 73-82

Schorin, C. (1992). *Modeling and projecting MBS prepayments, in Handbook of Mortgage Backed Securities*, edited by Frank J. Fabozzi, Chicago: Probus Publishing, 221-262.

Schwartz, E. & Torous, W. (1993). “Mortgage Prepayment and Default Decisions: A Poisson Regression Appoach.” *Journal of Real Estate Economics*, 21(4), 431-449

Stanton, R. (1995). “Rational Prepayment and the Valuation of Mortgage-Backed Securities.” *Review of Financial Studies*, Fall 8(3), 677-708

Wawro, G. (2006) “Advanced Topics in Maximum Likelihood Models for Panel and Time-Series Cross-Section Data.” Lecture note for 2006 ICPSR Summer Program