AN EMPIRICAL STUDY ON K-MBS PREPAYMENT PATTERNS

By

KIM, Min

THESIS

Submitted to

KDI School of Public Policy and Management

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF DEVELOPMENT POLICY

2018

AN EMPIRICAL STUDY ON K-MBS PREPAYMENT PATTERNS

By

KIM, Min

THESIS

Submitted to

KDI School of Public Policy and Management

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF DEVELOPMENT POLICY

2018

Professor Man CHO

AN EMPIRICAL STUDY ON K-MBS PREPAYMENT PATTERNS

By

KIM, Min

THESIS

Submitted to

KDI School of Public Policy and Management

In Partial Fulfillment of the Requirements

For the Degree of

MASTER OF DEVELOPMENT POLICY

Committee in charge:

Professor Man CHO, Supervisor ________________

Professor Min HWANG ________________

Professor Sherzod SHADIKHODJAEV ________________

Approval as of December, 2018

ABSTRACT

AN EMPIRICAL STUDY ON K-MBS PREPAYMENT PATTERNS

By

Min Kim

This study aims to analyze the determinants of prepayment risk in K-MBSs (Korean Mortgage Backed Securities) issued by KHFC (Korea Housing Finance Corporate). In total data used, 205 K-MBSs issued from 2004 to 2016 backed by 1,752,299 underlying assets that are all FRMs (Fixed Rate Mortgages). 14,573 observations (or month-security combinations) were recorded.

This research assumes that the five main causes of the risk are refinance incentive caused by the change of interest rate, cash-out refinance, housing turnover, curtailment, and default that the K-MBS backed by FRMs has very low default risk. Four independent variables, market (interest rate) spread, change rate of apartment purchase price, aging in month, and seasonal dummy, are estimated for the effects on latest 12-month CPR (Conditional Prepayment Rate).

Furthermore the study tries to control unobserved factors and the unbalanced panel of the study that is possible to have time heterogeneity caused by different issuance dates. In terms of methodology, a panel regression Fixed Effects model and a Two-way Fixed Effects model with month dummy and year dummy variables are considered. The study also assumes that Korean MBS market is lemon market caused by the asymmetric information. It can be considered as unobserved historical factors on dependent variable and independent variables. Therefore the study employs one and two-month lagged dependent variables as proxy independent variables and one and two-month lag effects on the two main explanatory variables, the market spread between the mortgage rate of K-MBS and the prime mortgage rate in market and the logarithm changing rate of apartment purchase price for controlling these unobserved factors.

Overall, all explanatory variables except seasonal dummy are significant. Based on the result of adjusted R^2, AIC, and F-test for H_0: $u_i = 0$, the most significant models are the 2nd order autoregressive (AR2) CPR model without lagged independent variables and the model with year dummy variables. The result summary of the AR2 CPR models is here,

1) When the market spread increases 1% point, CPR increases 0.57 ~ 1.07% point.

2) When ∆HPI of Apt. increases 1% point, CPR increases 0.4433 ~ 0.6405% point.

3) AR2 CPR model implies that 1% point increase in CPR one month ago affects 1.4519 ~ 1.5375% point increase in the CPR this year and that of CPR two months ago leads to an estimated 0.4732 ~ 0.5579% point drop in the CPR this year.

The time lag effects on CPR are similar in all six AR1 (1st order autoregressive) CPR models and all six AR2 CPR models regardless of F-test so that the time lag effects on CPR exist in this research. It is also considered that there might be dispersion of the time lag effects on CPR and the two main explanatory variables even though the models are jointly significant in p-value of F-statistic. Arellano-Bond estimation of 'Dynamic panel regression models' is applied to the optimal model for controlling these unobserved historical factors. Comparing with the result of non-dynamic panel regression model based on the study assumption, one and two-month lagged CPR show the similar effects so that the study assumption is reasonable.

Consequently, the pool-level of this study with some controlling measures about the unobserved factors in panel dataset has an advantage that the study method estimates easily and directly the main four variables on CPR. This study will enhance the understanding of the prepayment factors of K-MBS and will be the basis for pricing K-MBS.

Dedicated to my parents

ACKNOWLEDGEMENTs

It is the deepest honor and pleasure to give my great sense of gratitude to my supervisor Dr. Man Cho, and my advisor Dr. Min Hwang.

My great mentor and honorable philosopher is Dr. Man Cho. His devotion to help me had been solely and mainly responsible for the completion of my work. His timely advice and scholarly approach have helped me in accomplishing this task.

I owe a deepest sense of gratitude also to Dr. Min Hwang. He gave me the awesome advices that my work could be well organized and corroborated.

I appreciate all of my parents' supports that they gave me all of them with great virtue of patience and their prayer for me.

I owe whatever I achieved in KDI school to professors and friends.

Thanks so much.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Ⅰ. Introduction

1. Study purpose

The MBS (Mortgage Backed Securities) market plays an important role in the bond market around the world and the MBS market is closely related to the housing finance market since MBS securitizes mortgages as underlying assets. The MBS market in Korea depends largely on KHFC (Korea Housing Finance Corporate) that issues all kinds of K-MBSs[1] securitized with mortgages in Korea. KHFC has expanded long-term FRMs (Fixed Rate Mortgages) with capital raised from the capital market through K-MBS issuance since Korean government established KHFC in March 2004 based on the Korea Housing Finance Corporation Act. Under this Act, KHFC has a high level of credibility that KHFC in the condition of prepayment and default risks is sponsored by the government. Therefore the role of KHFC contributes to the development of Korean housing market as improving its intermediation efficiency as the financial inclusion for low and middle income households.

Although the KHFC's FRM MBS has very low default risk, mortgagors have the right to prepay their mortgages at any time before the maturities of mortgages. As economic subjects, they determine prepayment for various economic and social reasons. The nature of prepayment that mortgagors choose as an economic entity depends on the benefits and costs generated from economic, unexpected, or combined reasons. Economic reasons considered refinance incentive directly affect prepayment decisions and many studies take them as main determinants. Unexpected, non-economic, reasons not directly related to prepayment are mainly caused by disposing of houses when mortgagors make occupational or academic movements.

Consequently, this research examined MBS prepayment models based on the theoretical and

[1] The original name of K-MBS (Korean Mortgage Backed Securities) is KHFCMB2004S-01 issued solely by KHFC. Hereafter it is called K-MBS.

empirical study of prepayment of FRMs and then a rational model for K-MBS of KHFC and prepayment factors constituting the model are selected. The model and factors are analyzed practically with the data obtained from KHFC DB (Data Base) and public financial institutes such as BOK (Bank of Korea), KOFIA (Korea Financial Investment Association), and KFSS (Korea Financial Supervisory Service). This study will enhance the understanding of the prepayment patterns of K-MBS and will be the basis for pricing K-MBS. This research studying K-MBS suitable for Korea will be the basis for activating MBS market in Korea. The study will be expected to promote sound development of the housing finance market in Korea.

2. Development of K-MBS

The mortgage market in Korea has grown rapidly in the 2000s since the financial crisis, so called Korean IMF crisis. Even though GFC (Global Finance Crisis) occurred after the collapse of Lehman Brothers in September 2008, this financial event didn't affect the growth of mortgage market in Korea. According to BOK, total OLB (Outstanding Loan Balance) in April 2018 (the end of the study observation) is KRW 583.879 trillion and it has been gradually increased from KRW 292.813 trillion since December 2007.

The MBS market, which securitizes mortgages as underlying assets, plays an important role in the bond market around the world. In Korea, KoMoCo (Korea Mortgage Corporation) was established in September 1999 and issued MBS from 2000 to 2003. As a joint venture with Ministry of Construction, Housing and Commercial Bank, Kookmin Bank, Korea Exchange Bank, and Samsung Life Insurance Co., KoMoCo did not fully exercise the securitization function of housing finance. The largest issue amount was KRW 1.378 trillion of 2000 by the time when KHFC was established in 2004. There were two main reasons of this limitation that

KoMoCo dealing with MBS was restrictive to the continuous expansion of housing finance through MBS issuance. At first, there were few long-term FRM products that securitization incentive was small compared with floating rate mortgages. Secondly, KoMoCo at that time had low trust level.

To solve these problems, Korean government established KHFC in March 2004 based on the Korea Housing Finance Corporation Act. KHFC's main task is to expand long-term FRMs with capital raised from the capital market through MBS issuance. That is, banks provide mortgagors with long-term FRMs in the primary market and then KHFC securitizes these mortgages in the secondary market so that the long-term bond market is vitalized and housing finance is promoted. KHFC also has a high level of credibility since KHFC guarantees the MBS issued by KHFC in accordance with the Korea Housing Finance Corporation Act that the loss incurred by the KHFC is preserved by the government. K-MBSs are highly exposed to prepayment risk under internal and external market conditions even though most of K-MBSs are securitized from FRMs (underlying assets) that have relatively very low default risk. KHFC is also trying to figure out both prepayment and default risk and reduce the risks.

3. Four fundamental prepayment factors

There are many reasons of prepayment that previous researchers suggested. One of these priceless studies, Schorin (1992, 1995)[1), 2)] proposed four major factors, Seasonality, Burnout, Aging, and Interest rate. The brief introduction of Schorin’s four prepayment factors is summarized in this study.

Seasonality

Mortgage prepayments show seasonal trends according to specific patterns in month. The study data also showed the seasonal pattern (discussed it in Chapter Ⅳ) that SMM (Single Month Mortality) in March and April of spring and October and November of autumn was rising. The pattern is triggered from the household movement caused by marriage, occupational, and academic reasons. It can be evidenced from the change of HPI (Housing Price Index)[2]. As Schorin mentioned in his study, this factor is just a trend like his say, “all else being equal”. That is, this effect on prepayment can be overwhelmed by the bigger effective factors like refinance incentive. This study applied the factor with the seasonal dummy explained more in Chapter Ⅳ.

Burnout[3]

It indicates that prepayment rates follow not only the market spread between the mortgage rate of MBS and the predominant mortgage rate in market but also the path tracked by the two mortgage rates being equal or being differential. This factor is similar to the pool factor of Schorin (1992) who made the variable as a numerical value from 0 to 1. More explanation will be discussed in Chapter Ⅳ with Vintage analysis. In the case of this study, except that KHFC declined the sales of ‘u Bogeumjari’ and ‘Conforming loan (Fixed rate)’ from the second half of 2013 to the first half of 2014, the CPRs (Conditional Prepayment Rates) of K-MBS Pools (hereafter K-MBSs) have been increased by October 2015 since December 2008 when the mortgage rate in market started smaller than WAIR (Weighted Average Interest Rate) of KHFC. This path-dependent prepayment means that the smartest mortgagors recognized this refinance chance and prepay their mortgages while the remaining mortgagors are unaware of the chance.

[2] This study uses the annual and monthly average of apartment purchase price index since the incomparable portion of underlying assets on K-MBSs is apartment.

[3] Until the end of the study observation period (April 2018), only five K-MBSs (K-MBS0401, K-MBS0404~07) out of 205 K-MBSs have been terminated in the middle of their maturity periods so that the burn out effect is not included in this study.

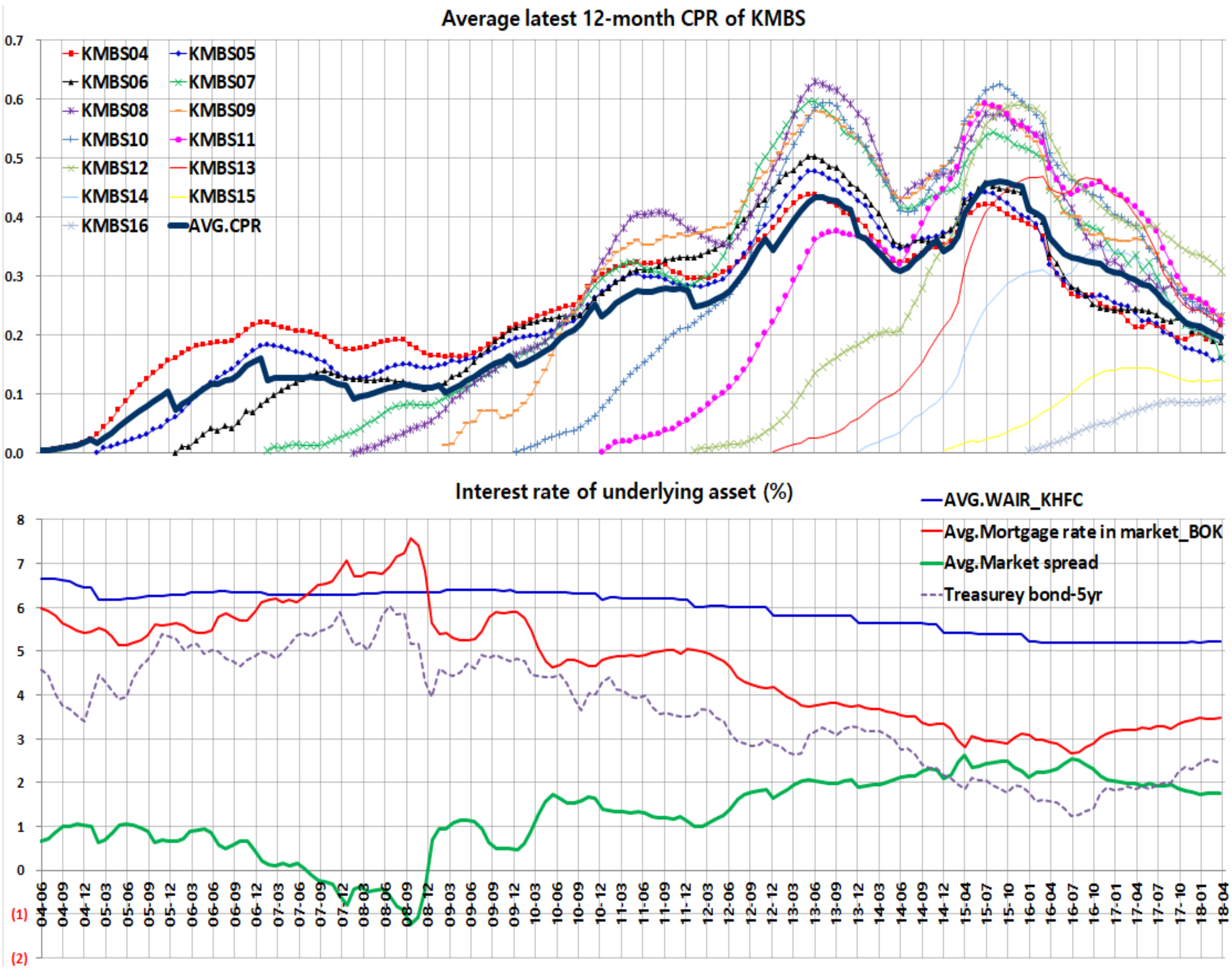
Aging[4]

Prepayment rates on MBSs have shown low level shortly since MBSs were issued and then gradually been augmented. Schorin pointed the effect with two causes fundamentally driven by the market spread between the mortgage rate of MBS and the predominant mortgage rate in market. First of all, the normal aging behavior of prepayment would occur unless the spread was expended enough to make refinance incentive soon after obtaining a new mortgage. The second example is related with the origination fee that mortgagors pay for a new mortgage. Basically, homeowners or mortgagors would not like to pay another origination cost when they take out a new mortgage except that the spread increased dramatically can surely result in refinance.

Aging effect in this study can be explained in Figure 1-1. Before December 2008 when the spread was stable and even from May 2006 to December 2008 when it was decreased, K-MBS04~07 showed the aging behavior since there was no outstanding refinance incentive originated from the movement of the spread like Schorin's explanation. After December 2008 when the mortgage rate in market was smaller than WAIR of KHFC, every K-MBS except K-MBS15, 16 didn't show the aging behavior. It can be said that there was obviously refinance incentive derived from the movement of the spread. The CPRs of all K-MBSs except K-MBS15, 16 started decreasing from October 2015 while K-MBS15, 16 showed the aging behavior since the spread started reducing from July 2016. For these reasons, K-MBS04~07 from May 2006 to December 2008 and K-MBS15, 16 from October 2015 showed the aging behavior while overall feature from December 2008 to October 2015 didn't show the aging behavior except the season when KHFC reduced the sales of 'u Bogeumjari' and 'Conforming loan (Fixed rate)' from the second half of 2013 to the first half of 2014.

[4] Aging in month is applied to the study.

Figure 1-1. CPR trends upon the path tracked by the market spread



Interest rate[5]

As mentioned continuously, there is no doubt that interest rate is the most fundamental factor on the prepayment of MBS. When KHFC makes their WAC (Weighted Average Coupon), KHFC uses the average of three five-year Treasury bonds from the past two days to issue date plus OAS (Option Adjusted Spread). As a proxy variable for this effect, this study employed the market spread between WAIR of KHFC and the primary mortgage rate in market obtained from

[5] This study applied at first the origination spread between WAIR and WAC of KHFC. However, there was no significant result so that it was replaced with the market spread between WAIR of KHFC and the primary mortgage rate in market obtained from BOK.

BOK. Schorin also mentioned that higher market mortgage rates (lower market spread) not only remove refinance incentive but also bring about a disincentive to mobility because of the movement cost including the disposal of existing houses and origination cost when a mortgagor takes out a new mortgage.

4. Controlling unobserved factors

This study tries to control unobserved factors such as curtailment that reduces the life of the loan shorter than the original payment schedule by paying off partially or totally the remained balance, loan characteristics, data quality (documentation level), and difference level (the smaller number of loans the fewer mortgagors) among K-MBS pools. The unbalanced panel of the study is possible to have time heterogeneity caused by different issuance dates. In terms of methodology, a panel regression Fixed Effects model and a Two-way Fixed Effects model with month dummy and year dummy variables are considered.

The study also assumes that Korean MBS market is lemon market caused by the asymmetric information between investors and mortgagors, smart investors and KHFC, or smart mortgagors and less smart mortgagors. The lemon market ex-ante concerns the willingness to pay the mortgages while the payment behavior like prepayment is revealed ex-post. It can be considered as unobserved historical factors on dependent variable such that relatively older loans have been well known to mortgagors and younger loans vice versa have effect on the dependent variable. The study employs one and two-month lagged dependent variables as proxy independent variables for controlling these unobserved factors.

In addition to the proxy variables, unobserved historical factors can also be considered on independent variables. Schorin also pointed out in 1995 study that the prepayment model must include a lag structure[6] between a change in mortgage rates and a prepayment behavior. There are smart mortgagors who can recognize the market situation and decide whether opening or closing a new mortgage and whether keeping or applying their decision to housing market. They can influence the other mortgagors. That is, time differential among mortgagors obviously exists like his analogy, 'pipeline lag' (transportation lag) in system operation. In this study, the two main explanatory variables, the market spread between the mortgage rate of K-MBS and the prime mortgage rate in market and the logarithm changing rate of apartment purchase price, are possible to correlate with omitted variables that estimating the variables is difficult. Therefore, one and two-month lag effects on the market spread and the change rate are applied.

5. Research design

A rational model for K-MBS of KHFC is established and selected prepayment factors are analyzed in this study. In order to directly estimate prepayment factors, one of pool-level linear models following the Schorin's method (1992, 1995) is chosen as a parsimonious model. The main data for description and inference analysis such as issue amount, total current balance, WAC, WAIR, SMM, and the information of underlying assets of K-MBS, etc. are obtained from KHFC and the market data such as primary mortgage rate, issue amount of ABS (Asset Based Securities), and HPI are acquired from other public financial institutes such as BOK, KOFIA, and KFSS. Nevertheless, there are some unobserved factors in the study model so that panel data analysis with FE (Fixed Effects method) is performed. The result of the research can be expected to help understand prepayment patterns of K-MBS and promote K-MBS market and Korean housing finance market. The composition of this study is as follows.

[6] This study applies one and two-month lag effects on the change of apartment purchase price in addition to the interest (mortgage) rates.

After the Korean mortgage-MBS market is introduced in Chapter Ⅱ, Chapter Ⅲ presents literature review based on models categorized into three types, pool-level linear model, pool-level nonlinear model, and loan-level model. In Chapter Ⅳ, description analysis and the information for inference analysis such as data resources, variables, and theories for making the study model are explained. The panel data set (total pool of FRM K-MBSs) is constructed with annual pools aggregated in each issue year from 2004 to 2016 and then regression analysis is performed with STATA (Statistics and data analysis tool). The study result is explained in Chapter Ⅴ. Based on the above analyses, implications for the development of the K-MBS market are derived. In particular, the study with suggested prepayment models proposes reasonable suggestions for promoting Korean housing finance market.

Ⅱ. The Korean mortgage-MBS market

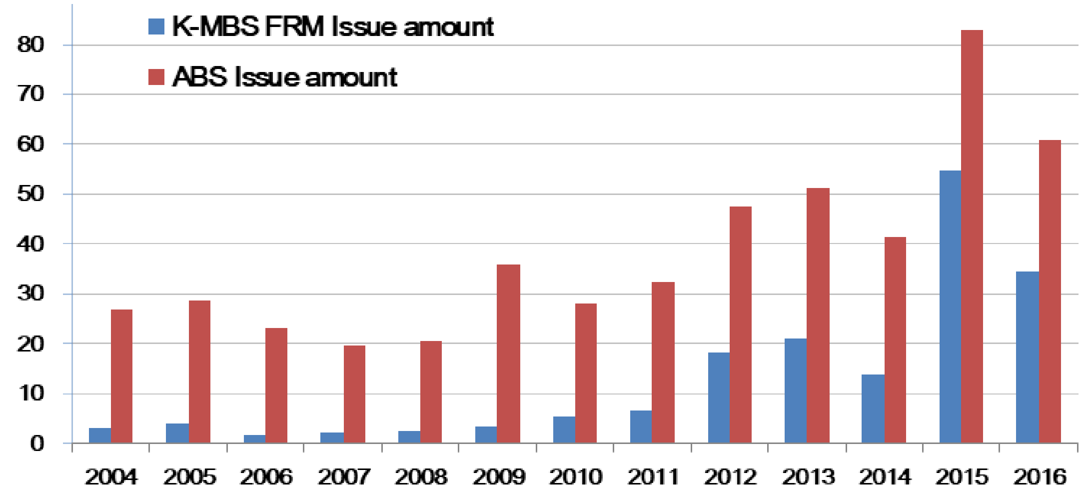
The whole types of underlying assets applied for K-MBS at a glance is shown in Table 2-1. Comparing the size of K-MBS issuance with that of ABS[7] in Figure 2-1, the issuance trend of K-MBS is similar to that of ABS (Overseas issuance included). Since 2004 when KHFC was established, the issuance size has been increased by 2015 that the total K-MBS was KRW 54.7723 trillion out of the total ABS, KRW 82.9082 trillion. Especially from 2012, the issuance has been significantly increased due to the rapid increase of the long-term FRMs which have a

Table 2-1. Mortgage types of K-MBSs

Loan Name (interest rate type) Securitizing years	Number of loans	Loan amount, KRW 1million (Ratio %)	Loan Name (interest rate type) Securitizing years	Number of loans	Loan amount, KRW 1million (Ratio %)
t/e Bogeumjari (Fixed) 2004~2013	284,384	23,216,733 (13.67)	Didimdol (Adjusted every 5yrs) 2014~2016	3,836	349,156 (0.21)
u Bogeumjari (Fixed) 2010~2016	460,899	47,458,795 (27.94)	Didimdol (Fixed) 2014~2016	132,203	12,641,017 (7.44)
Conforming (Adjusted every 5yrs) 2014~2016	237,946	24,624,631 (14.50)	t+ Bogeumjari 2015~2016	15,640	1,885,157 (1.11)
Conforming (Fixed) 2012~2016	288,083	28,010,528 (16.49)	Ansim Conversion (Adjusted every 5yrs) 2015	17,073	1,786,250 (1.05)
Conforming (Mid-term fixed) 2014~2016	2,789	186,608 (0.11)	Ansim Conversion (Fixed) 2015	309,420	29,703,510 (17.49)
Conforming (Debt adjustment fixed) 2014, 2016	26	2,128 (0.001)	Total 2004~2016	1,752,299	169,864,514 (100)
Note. - Below KRW 1million, loan amount was cut					

[7] KFSS (Korea Financial Supervisory Service) evaluates the Korean ABS market annually and the report is shared to the public.

Figure 2-1. Annual total issue amount of K-MBS vs. ABS, KRW 1 trillion



relatively low interest rate like ‘u Bogeumjari’ of KHFC and ‘Conforming loan (Fixed rate)’ of financial institutes. Conforming loan issued first in March 2012 is a mortgage suitable for securitization that meets the loan conditions of KHFC. It was driven by ‘Mortgage policy of the government, 2011.06.28’ to stabilize the household debt and the policy guided that domestic banks increase the proportion of long-term FRMs by 30% until 2016.

According to the annual ABS analysis report 2012 of KFSS, the size of securitization based on the mortgage bonds of KHFC was KRW 20.2813 trillion in 2012, which is higher than the amount of 2011, KRW 10.0602 trillion. The increased amount of K-MBS 2012 is KRW 10.22 trillion, 101.6% higher than that of K-MBS 2011. It exceeded the average amount of K-MBS KRW 6.3 trillion issued between 2004 and 2011. The amount of K-MBS was decreased in 2014 since KHFC declined the sales of ‘u Bogeumjari’ and ‘Conforming loan (Fixed rate)’ from the second half of 2013 to the first half of 2014. The annual ABS analysis report 2014 of KFSS explained that it was caused by expanding the market spread[8] from 2012 (Figure 1-1). Even

---

[8] Market spread in this study means that the spread between the mortgage rates based on KHFC’s WAIR and those based on financial institutes.

though the expansion of the spread kept from 2012 to the first half of 2016, the largest amount of K-MBS was recorded in 2015.

According to the annual ABS analysis report 2015 (KFSS), the securitization size based on the mortgage bonds of KHFC was KRW 55.2 trillion in 2015, which is significantly higher than the amount of 2014, KRW 14.5 trillion; K-MBS issuance increased by KRW 40.7 trillion, (+ 280.5%) compared to the previous year 2014. Table 2-1 represents that the soaring increase was caused by 'Ansim Conversion loan'. Based on the annual report 2015, the increase was also caused by the recovery of the housing market that made the existing Jeonse (Korean specific housing charter) demand converted to real purchase demand in the housing market.

K-MBSs issued from 2004 to 2013 were composed of one type of loan, 't/e Bogeumjari' or 'u Bogeumjari' for each K-MBS. 'Conforming loans' with fixed rate (so called 'basic type') from 2012 to 2013 were also pooled for one K-MBS. On the other hand, K-MBSs issued since 2014 have been pooled with several types of loans. For example, a K-MBS named as KHFC MB2015S-17 was pooled with 't+ Bogeumjari', 'u Bogeumjari', 'Didimdol ' (adjusted rate every five years, fixed rate), 'Conforming' (adjusted rate every five years, fixed rate, and mod-term fixed rate), 'Ansim conversion' (adjusted rate every five years, fixed rate).

By these mortgage policies, governments support low and middle income households for housing stability and the government policy can also influence the prepayment decision. For instance, Korean government employed 'Ansim Conversion loan'[9] in 2015 for improving the household debt structure. This government role in finance sector had effect on the highest prepayment rate in 2015 out of K-MBS history since 2004. Korean major housing finance policies of mortgages for low and mid-income households are presented in Table 2-2[3)]. Cho

[9] Ansim Conversion loan: A bank loan product that converts short-term, floating interest rate, and interest only loan products into long-term, fixed interest rate, and amortization loan products.

(2012)[4] also suggested policy implications for Korea mortgage market based on two aspects, financial stability and financial inclusion. It is introduced in Table 2-3.

Table 2-2. Housing policies of mortgages for low and mid-income households (Continue)

Loan name	House price limit, KRW 1million	* Interest type (Rate, %)	Maturity (Grace), year	Repayment type (Prepayment penalty, %)
Didimdol	600	Fixed, Adjusted every 5yrs (2.3~3.1)	10, 15, 20, 30 (1 or no grace)	CAM, CPM (Within 3yrs, 1.2 in * sliding)
Profit sharing type	600 (only apartment)	Fixed (1.5)	20 (1 or 3)	CAM, CPM (Within 3yrs, 0.9~1.8)
Stable housing for purchase	600	Adjusted (3.3)	20 (1 or 3) 30 (5)	CAM, CPM (No penalty)
Bogeumjari	900	Fixed (2.85~3.1, '15.04')	10, 15, 20, 30 (1 or no grace)	CAM(incremental), CPM (No penalty)
Conforming	900	Fixed (2.80~4.15)	10, 15, 20, 30 (1 or no grace)	N.A. (1.5)
Beotimmog Jeonse	Deposit 200 (Capital area 300)	Adjusted (2.7~3.3)	2~10, 4times extension (No grace)	IOL (No penalty)
* Low-income household Jeonse	N.A.	Adjusted (2.0)	12 (No grace)	CAM, CPM, Mixed (No penalty)
Stable housing for monthly rent	Deposit 100, Rental 0.6	Adjusted (2.0)	3~6, 3times extension on a year (No grace)	IOL (No penalty)

* Note.
- Prime rate (% point discount) of Didimdol, Stable housing for purchase, and Beotimmog Jeonse loan: Household with three or more children (0.5), Multi-cultural family (0.2), Family with a disabled member (0.2), and First home buyer (0.2). Only one discount can be applied
- Sliding method: Decrease applied within designated period from the loan origination depending on the number of remaining days, n/365
- Interest rates of Stable housing for purchase and Low-income household Jeonse are operated by 'National Housing Fund Plan'
- CAM : Constant Amortization Mortgage, CPM: Constant Payment Mortgage
- IOL : Interest Only Loan

Table 2-2. Housing policies of mortgages for low and mid-income households

Loan name	Non-homeowner	* Income limit (First home buyer) KRW 1million	Loan limit KRW 1million	Region limit
Didimdol	○	60 (70)	Up to 200	N.A.
Profit sharing type	5yrs more (First home buyer)	60 (70)	Up to 70% of house value (Up to 4.5 times of * Income limit)	Capital area, Metropolitan cities (Population over 500,000), Sejong city
Stable housing for purchase	○	60	Up to 200	N.A.
Bogeumjari	○	70 (Newlyweds 85)	Up to 70% of collateralized house value, up to 500	N.A.
Conforming	N.A.	N.A.	Up to 500	N.A.
Beotimmog Jeonse	○	50 (Newlyweds 55)	Up to 70% of deposit, up to 80	Capital area, KRW 100 million
Low income household Jeonse	○	N.A.	Up to 70% of deposit	N.A.
Stable housing for monthly rent	○	Up to 30 (parents income base)	N.A.	N.A.

* Note.
- Income limit: Combined annual income of the head of a household and his/her spouse
- Housing pension: Older than 60years, Up to KRW 900 million of house value (In the case of homeowners, up to KRW 900 million, the existing house included)

Table 2-3. Policy implications for Korean mortgage market

	Financial stability	Financial inclusion
Mortgage	☐ Ratio of FRM (e.g. 30% of MDOs) ↑ ; Funding through MBS, CB ↑ ☐ Set interest rate cap on ARM ☐ Ratio of Long-term(10~30yrs) CAM ↑ ☐ Ratio of hybrid of 5 and 10yrs FRM ↑	☐ Exempting prepayment penalty on ARM ☐ Setting mortgagor's financial crisis provision on FRM ☐ Extending adjustment cycle of interest rate on ARM (1yr level) ☐ Studying non-standard mortgage product (e.g. VMM, GPM, Buy-Back)
Examination	☐ Maximizing LTV, DTI as a Market Stabilization (e.g. LTV<80, DTI<50) ☐ Regulating prime and sub-prime market: government role in each sub-market ☐ Monitoring loan purpose, multiple house owner’s loan: Consider compensating Risk	☐ Activating mortgage insurance market (through introduction of public mortgage insurance) ☐ Enhancing financial service for marginal mortgagor through continuous expansion of prime market ☐ Targeting specific consumer (self-employed, 20’s~40’s low-income mortgagor, first home buyer, etc.)
Funding	☐ Ratio of funding through MBS, CB ↑ (e.g. 20~30% of MDOs) ☐ Regulating stable underlying asset provision for MBS and CB pool (e.g. US QRM) ☐ Studying MBS, CB product suitable for Korea ☐ Monitoring unnecessary complexity of product	☐ Linking wholesale funding and product and assessment criteria for marginal mortgagor (enhancing mortgage liquidity of marginal mortgagor) ☐ Establishing prime MBS market and regulating government role (guarantee etc.) ☐ Studying consumer-friendly funding scheme like Danish mortgage bond market
Regulation	☐ Comparing and analyzing various regulatory measures including LTV and DTI regulation (Differentiating weighted risk, establishing allowance for bad debt etc.) ☐ Developing long-term and continuous regulatory measure (minimizing market uncertainty)	☐ Considering support policy such like tax credit for marginal mortgagor ☐ Analyzing effectiveness of financial support policy for low-income mortgagor like Jeonse loan (overseas case analysis like Taiwan)
Note. - FRM: Fixed Rate Mortgage - ARM: Adjusted Rate Mortgage - MDO: Mortgage Debt Obligation - MBS: Mortgage Backed Security - CB: Covered Bond - CAM : Constant Amortization Mortgage - VMM: Vanguard prime Money Market - GPM: Graduated Payment Mortgage - LTV: Loan To Value - DTI: Debt To Income - QRM: Qualified Residential Mortgage - Jeonse: Korean specific housing charter		

Ⅲ. Literature review

This study fundamentally applied the prepayment model proposed by Schorin (1992, 1995) since this study tried to suggest a parsimonious model to directly estimate prepayment factors. Especially, Schorin proposed significant explanatory variables, interest rate, burnout, seasonality, and aging on prepayment, CPR. His model can be categorized into the pool-level linear model. This study summarized a literature review based on models categorized into three types, pool-level linear model, pool-level nonlinear model, and loan-level model. These three types are also called the economic model of prepayment determinants. After this study introduces 'Standard prepayment benchmark model' (hereafter PSA standard) that PSA (Public Securities Association) traditionally has announced the standard, the three types of the economic model are summarized.

1. PSA standard

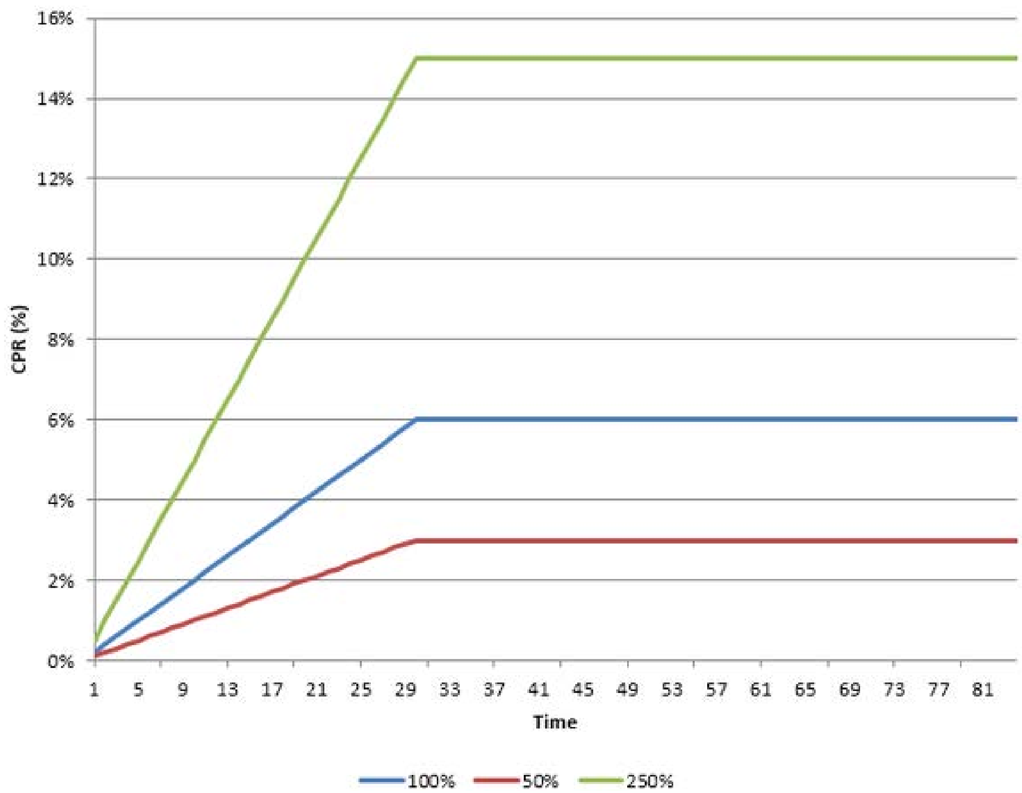
The standard uses prepayment data of the MBSs issued by GNMA(Government National Mortgage Association, Ginnie-mae), FNMA(Federal National Mortgage Association, Fannie-mae), FHLMC(Federal Home Loan Mortgage Corporation, Freddie-mac). PSA proposed CPR in the next equation as the prepayment benchmark model and the feature is shown in Figure 3-1.

$$CPR^{10} = \frac{6}{30}t \text{ , if } t \leq 30 \text{ and } 6 \text{ if } t > 30$$

PSA standard can be used for the rough analysis of the mortgages, MBS cash flows, and pricing MBSs since it depicts simply but accurately the relationship between the prepayment

Figure 3-1. PSA standard

[10] CPR is annualized with SMM in this equation.
$CPR = 1- (1-SMM)^{12}$; $(1-SMM)^{12}$ means that it is not prepaid until 12 months.



and the time course of a typical mortgage pool. However, it is not used to predict the prepayment of a particular mortgage and mortgage pool in terms of relying on simple historical experience. PSA standard is not taking into account various economic variables such as the interest rate and the housing price related to the prepayment decision of the economic entity.[5)]

2. Pool-level linear model

This model type is the initial empirical model in comparison with other types. The representative models are Asay, Guilaume, and Mattu (1987)[6)], Brazil (1988)[7)], Carron and Hogan (1988)[8)], Davidson, Herskovits and Van Drunen (1988)[9)], Chinloy (1991)[10)], Schorin (1992, 1995) and others. The basic analysis subjects of these models are not individual mortgages but mortgage pools aggregated with several mortgages. In this study, Asay, Guilaume, and Mattu (1987), Chinloy (1991), and Schorin (1992) are explained. Since these models have a simple structure like Equation (3-1) and the required data for estimation are relatively little, it has an advantage that efficiently measures the prepayment rate.

$$CPR_{i,t} = \alpha + \beta \square Spread_{i,t} + \sum_k \Gamma_k \square X_{k\ i,t} + \varepsilon_{i,t} \quad (3\text{-}1)$$

Asay, Guillaume, and Mattu (1987)

presented a linear model using only ‘Spread term’ between the contract interest rate and the market mortgage rate in ‘arctangent’ function as an explanatory variable. They collected the data in time-series without identity. Fixed additional interest rate considering the maturity of mortgages was added in the ‘arctangent’ function and error term ε_t means stochastic noise term.

By adjusting the parameters appropriately, they modeled precisely the phenomenon that the prepayment rate starts to rapidly increase at a certain point and gradually stops. According to the model, the prepayment rate gradually increased at a certain level and then the growth rate slowed down in the condition that the contract interest rate is lowered and the market mortgage rate becomes relatively higher. However, the explanatory variable is the only one, difference between the contract interest rate and the current market mortgage rate, so that they couldn’t explain other aspects of the prepayment.

Chinloy (1991)

estimated the prepayment rate with a linear model considering the explanatory variables, contract interest rate, market interest rate and the age of mortgages. The analysis period recording the data in month was from 1986 to 1989, and the model was estimated with the whole mortgages of

GNNA's MBS. He just used contract and market interest rates of mortgage i at time 't' as for 'Spread term' and age of mortgage in month was added for considering another variable.

According to the study result, the lower the current market interest rate and the higher the contract interest rate, the higher the prepayment rate when all other conditions are the same. However, the result showed that the explanatory level of variables for reflecting the degree of maturity was weak and it resulted from the simple assumption that the prepayment rate linearly increases as the maturity goes on. Although Chinloy model is a more advanced than that of Asay et al.(1987) in the aspect of considering the maturity, the model has limitation that the two interest rates were used for each independent variable instead of reflecting the spread between the contract interest rate and the current market interest rate.

Schorin (1992, 1995)

analyzed the prepayment rate by pooling annually the GNMA's 30-year MBS mortgages from 1973 to 1989. He estimated the prepayment rate with the ratio[11] between the contract interest rate and the current market mortgage rate, age, seasonal movement dummy, and the WA (weighted average) outstanding loan balance of annual mortgage pools.

In addition to these explanatory variables, macroeconomic variables such as industrial production, the amount of new mortgages issued in month, new housing construction permits and housing sales in month, interest changing rate in three months, and unemployment rate were added to analyze the prepayment rate, however his additional variables didn't show any significant effects on the prepayment. The main variables are employed in this study and the study models with the main variables will be explained in Chapter Ⅳ.

[11] $PSP_{i,t} = (rc_{i,t} - rt_{i,t}) / rc_{i,t}$; The ratio of quantified refinance incentive.
- $rc_{i,t}$: Contract interest rate of mortgage i at time 't'
- $rt_{i,t}$: Market interest rate of mortgage i at time 't'

3. Pool-level nonlinear model

This model uses a mortgage pool as a base unit and the model has very nonlinear characteristics. Richard and Roll (1989)[11)] developed one of the most well-known nonlinear prepayment model commissioned from Goldman Sachs. They argue that the prepayment rate at specific time points is determined by the interaction among the four major explanatory factors, refinance incentive, seasoning (aging), seasonality, and burnout explaining the volatility of prepayment. His model is shown in Equation (3-2).

$$CPR_t = RI_t \square SM_t \square MM_t \square BM_t \quad (3\text{-}2)$$

- CPR_t : Conditional Prepayment Rate occurred at time 't'

- RI_t : $\alpha + \beta \square \arctan(\Gamma \square SP_t + \delta)$ of Asay, Guillaume, and Mattu (1987)

- Refinance incentive at time 't' is originated from Asay, Guillaume, and Mattu (1987) model and estimated with NLLS (Non-Linear Least Squares)

- SM_t : Seasoning multiplier at time 't' = $\frac{1}{30} a_t \square I[at \leq 30] + I[a_t > 30]$; 'I' is indicator of the factor

- Based on PSA standard (i.e. CPR = $\frac{6}{30}t$, if $t \leq 30$ and 6 if $t > 30$), this factor estimates the value that linearly increases by $\frac{1}{30}$ every month until 30 months and become constant after 30 months

- MM_t : Seasonality(month) multiplier at time 't' = $\zeta \square \sin(\eta \square m_t + \theta)$

- Prepayments generally become high in the later summer and low in the winter since households are moving in this seasonal pattern. A sine curve represents the factor

- BM_t : Burnout Multiplier at time 't' = $\exp[k\{\prod_{i=1}^{t} \max(r_c / r_i , 1)\}]$

- It means that some smarter mortgagors recognized the refinance incentive and prepay their mortgages while the remaining mortgagors are unaware of the incentive

Many nonlinear models derived from the Richard and Roll model have been proposed. The modified Goldman Sachs model and the OTS (Office of Thrift Supervision) model are the most widely accepted models. The models are so nonlinear that the estimated parameters are not stable. Nevertheless, these models attempted to make a relatively sophisticated estimation using significant factors based on the prepayment theory.

4. Loan-level model

This analysis model is a relatively recent prepayment model in the pace of accessibility to and statistical development of individual mortgage data. This model types take into account the characteristics of the mortgagor and the nature of the mortgage contract directly and specifically. The limited dependent variable method that considers individual prepayment determinant features and the duration analysis that fully takes into account the information included during the time of prepayment are used for this model. Some models consider prepayment and default determinations not as independent decisions but as competing risks. As this research that the KHFC's FRM MBSs are securitized from their underlying assets with very low default risk, some other models considers prepayment risk without default risk. The case of OFHEO (Office of Federal Housing Enterprise Oversight) that used simple logit model instead of MNL (Multi-Nomial Logit) model for the prepayment model, and the case of Schwartz and Torous (1989) that used Cox PHM (Proportional Hazard Model) are the examples of some other models.

OFHEO[12)]

invented a prepayment model in order to ensure the capital adequacy and financial safety and soundness of two housing government-sponsored enterprises (GSEs), Fannie Mae and Freddie Mac. The prepayment model used individual mortgage data and estimated the prepayment rate by considering the prepayment and default of mortgages simultaneously according to interest types. The model based on MNL is next and the data of Equation (3-6) obtained in month are various such as LTV (Loan to Value) ratio of mortgage i at Origination, PNEQ (Probability of Negative Equity) on the principal of mortgage i at time 't', burned out effect of mortgage i at time 't', and relative loan size (weighted average) mortgage i at Origination. More information of variables can be obtained from the footnote of Equation (3-6).

$$Pr(Y=Prepayment|X) = \frac{\exp(X\square\beta prepayment)}{1+\exp(X\square\beta prepayment)+\exp(X\square\beta default)} \quad (3\text{-}3)$$

$$Pr(Y=Default|X) = \frac{\exp(X\square\beta default)}{1+\exp(X\square\beta prepayment)+\exp(X\square\beta default)} \quad (3\text{-}4)$$

$$Pr(Y=Current|X) = 1 - Pr(Y=Default|X) - Pr(Y=Prepayment|X) \quad (3\text{-}5)$$

$$Y_{i,t} = \sum_k \beta_k \square X_{k\,i,t} + \varepsilon_{i,t} \quad (3\text{-}6)$$ [12]

5. Others

Schwartz and Torous (1989)[13)]

explained the conditional probability function of prepayment with a proportional hazard function according to Cox's method (1972, 1975)[14), 15)]. In this estimation process, pool-level data aggregated with mortgages rather than individual mortgages were used in their model, Cox PHM

[12] More details are in Title 12 of the Code of Federal Regulations of U.S. and https://www.law.cornell.edu/cfr/text/12/appendix-A_to_subpart_B_of_part_1750
OFHEO
https://web.archive.org/web/20061102175242/http://www.ofheo.gov/

(Proportional Hazard Model). Cox PHM is a kind of the duration analysis that controls variables shown their effects during the time of prepayment. Although they applied only the age, refinancing rate resulting from interest rate difference, the acceleration effect of the refinancing rate, burnout effect, and seasonality of mortgages into the model, their Cox PHM as a probability function of prepayment has the biggest advantage that time-varying covariates employed in the model contain the cost of refinancing, demographic variables characterizing mortgagors, and geographic factors particular to local mortgage markets. The Cox PHM modeled for the prepayment function is shown in Equation (3-7).

$$\pi(t;\nu',\theta') = \pi_0(t;\Gamma,P)\square\exp(\beta'\nu')\ ;\ \nu' = 0 \text{ where homogeneous condition} \quad (3\text{-}7)$$ [13]

- $\beta' = (\beta1, \beta2, \ldots, \beta s)$: a (column) vector of regression coefficients

- $\nu' = (\nu_1, \nu_2, \ldots, \nu_s)$: a (row) vector of time-dependent (-varying) explanatory variables or time-varying covariates until prepayment

- $\theta' = (\theta_1, \theta_2, \ldots, \theta_k)$: a vector of parameters to be estimated

$\pi_0(t;\Gamma,P)$[14] is the base line hazard function given by the log-logistic hazard function and it is shown in Equation (3-8); 'Γ' is scale parameter and 'P' is shape parameter

$$\pi_0(t;\Gamma,P) = \Gamma\square P\ \square(\Gamma\ \square t)^{p-1} / \{1+(\Gamma\ \square t)^p\} \quad (3\text{-}8)$$ [15]

HR(Hazard Ratio)[16), 17)] can represent the prepayment rate and it is shown in Equation (3-9).

---

[13] $\exp(\beta'\nu') = \lim_{t\to\infty}\left(1+\frac{\beta'\nu'}{t}\right)^t = \sum_{t=0}^{\infty}\beta'\nu'/t!$

[14] It measures by itself the probability of prepayment where homogeneous condition, $\nu' = 0$

[15] If $P>1$, the probability of prepayment increases from $\pi_0(0) = 0$ to a maximum at $t^* = (P-1)^{1/P} / \Gamma$ and then decreases to $\lim_{t\to\infty}\pi_0(t) = 0$

$$HR = \pi_0(t) \square \exp\{\sum_{k=1}^{j-1} \beta_k \nu_k + \beta_j(\nu_j + c)\} / \pi_0(t) \square \exp(\sum_{k=1}^{j-1} \beta_k \nu_k + \beta_j \nu_j) = \exp(c\beta_j) \quad (3\text{-}9)$$

Where ν' is a continuous scale covariate vector and a constant $c = 1$,

- HR=1 ($\beta_j = 0$) j^{th} explanatory variable has no effect on the event
- HR>1 ($\beta_j > 0$) j^{th} explanatory variable increases hazard effect on the event
- HR<1 ($\beta_j < 0$) j^{th} explanatory variable decreases hazard effect on the event

The explanatory variables have the same proportional effect at all mortgage ages and the vector of covariates (ν') means that independent variables and other noise factors share their variations. Therefore, ν' that has a higher prepayment rate at a particular mortgage age causes a higher prepayment rate at any other mortgage age. As a result, the vector of regression coefficients (β') can measure the effect of explanatory variables or covariates upon the prepayment determination.

Ⅳ. Data and variables

1. Data resources

KHFC provides a lot of useful information about K-MBS issued by KHFC through K-MBS DB[16] and Bloomberg Terminal. Since 2014, KHFC has provided 'CLC document' that researchers can acquire data about underlying assets of K-MBSs such as the payment histories of LTV and geographic distribution, latest 12-month CPR history, and delinquency history.

However KHFC pointed out that the utilization level of the DB is somewhat restrictive in making various decision processes although KHFC has a plenty of its internal data related to its services.[18)] As mentioned in its report, micro information DB of K-MBS including housing information is not completely established so that more specific analyses linked with the housing market is limited. Nevertheless, this study uses 'Base Asset Detail Information' provided by KHFC DB to analyze the prepayment factors through Panel data analysis with FE.

Especially in this study, 'Loan Level Data[17]' and 'MBS Cashflow Information' about the K-MBS backed by FRMs were mainly used to make panel dataset. The interface of 'Loan Level Data' and that of 'MBS Cashflow Information' are introduced in APPENDIX and more supplement information as loan-level data about underlying assets can be referred from 'Bloomberg Investor Report' and 'CLC document'.

2. Data description

This study analyzes totally 205 FRM K-MBSs issued from 2004 to 2016 and securitized with total 1,752,299 underlying assets. 14,573 observations were recorded monthly from June 2004 to April 2018 by KHFC. Its descriptive statistics for the entire data set is shown in Table 4-1. More details of the annual pools are explained in APPENDIX. The overall average issue amount of K-MBS is KRW 835,506 million from the total issue amount KRW 171,278,883

---

[16] K-MBS is the name of KHFC DB
http://kmbs.hf.go.kr
[17] Since the characteristics are close to pool-level data, it is hard to obtain the loan-level data.

million. The smallest issue amount is KRW 210,600 million and the biggest one is KRW 5,062,800 million. Total current LB (Loan Balance) as the UPB (Unpaid Principal Balance) by April 2018 is totally KRW 70,277,948 million. It is about 41% of the total issue amount and 59% of the amount has been already paid off or paid to the investors. The average of initial WAIR is 4.49% and that of current WAIR by the end of observation (2018.04) is 4.31%. The average of WAC is 3.55%, the origination spread[18] (initial WAIR – WAC) is 0.94, SMM is 3.08%, and the latest-12-month CPR is 0.28.

The result of ‘Vintage analysis in monthly aging’ in Figure 4-1 is to show seasoning effect of K-MBS and it was measured by the result of Accumulation amount of prepayment over Total initial LB. This method is similar to measure the pool factor[19] of Schorin (1992) who made the variable as a numerical value from 0 to 1 to observe the burnout effect of MBS. This measure represents that the prepayment speeds of K-MBS08 ~ 14 are higher rising than the speeds of K-MBS04 ~ 07. Until the result reached 60%, the prepayment speeds of K-MBS08 ~ 14 took about 32 to 43 months while those of K-MBS04 ~ 07 took about 58 to 66 months.

The main reason of this difference as shown in Figure 4-1 is that the mortgage interest rate in market declined after GFC (Global Finance Crisis). After the collapse of Lehman Brothers in September 2008, the expansion of the market spread started from December 2008 (Figure 1-1) and it made mortgagors refinance incentive. Especially, the change rate of the spread between 2008 and 2009 is the largest one so that the prepayment speed of K-MBS09 was faster than any other pools. In addition, the increase of K-MBS led by the increase of ‘u Bogeumjari’ since

Table 4-1. Descriptive statistics of 205 K-MBS pools

[18] Regression and Panel data analysis showed that the spread has no significant effect on CPR.
[19] Pool factor = outstanding principal balance / initial loan amount.
The pool factor means that the closer to 0 than expectation the value is, the faster the pool factor is declining than expectation and the more prepayment is made in the original payment schedule.

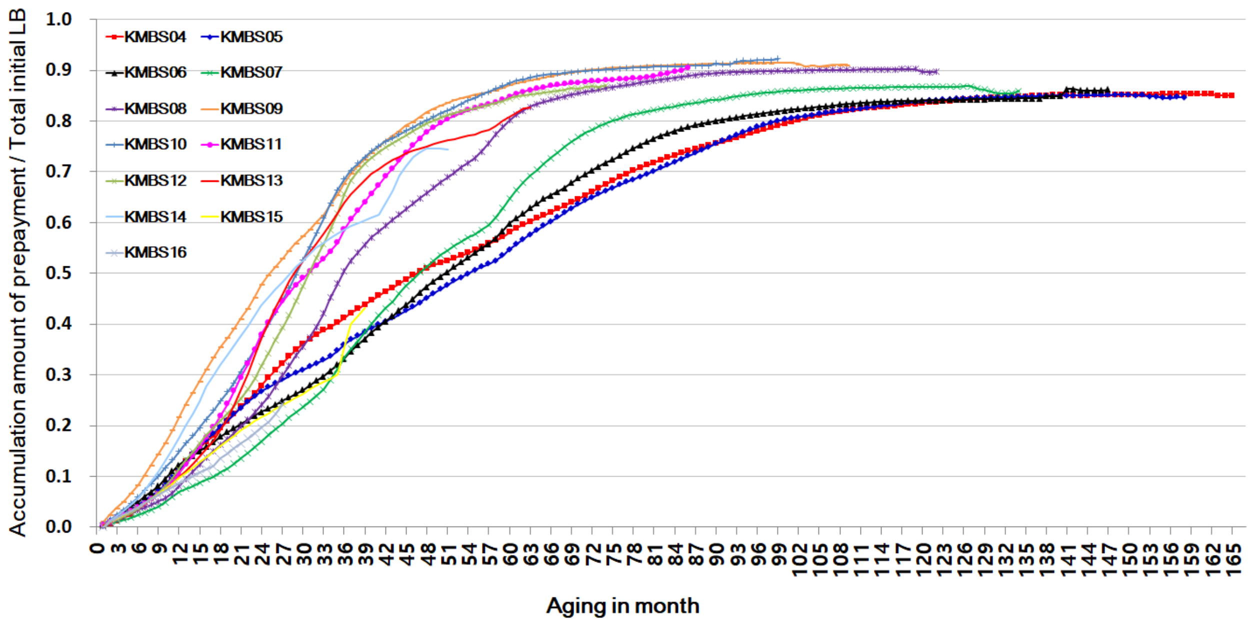
	Population Mean	Population SD	Population Min	Population Max
Issue amount (KRW 1million)	835,506	790,653	210,600	5,062,800
Total current LB, 2018.04 (KRW 1million)	342,819	623,185	0	2,888,629
WARM (month)	175.1	66.1	0.0	291.0
Number of loans	8,547	8,110	2,198	51,441
Number of pools	15	9	5	33
Initial WAIR (%)	4.49	1.33	2.28	7.22
Current WAIR, 2018.04 (%)	4.31	1.49	0.00	7.34
WAC (%)	3.55	1.23	1.51	6.64
Origination spread (bp)	94	42	-1	227
Avg. CPR (%)	28.32	17.03	0.00	73.56
Avg. SMM (%)	3.08	2.45	0.00	23.12

Note.
- Below KRW 1million, issue amount and total loan balance were cut
- LB, Loan Balance
- WARM, Weighted Average Residual Maturity
- WAIR, Weighted Average Interest Rate
: Average interest rate of the underlying pool's asset, weighted by current month’s loan balance
- WAC, Weighted Average Coupon
: The average of three 5-year Treasury bonds from the past 2dyas to issue date + OAS (Option Adjusted Spread)
- Origination spread = Initial WAIR of KHFC – WAC of KHFC
- CPR, Latest 12-month Conditional Prepayment Rate = $\{1 - \prod_{k=0}^{11}(1 - smm_{12\text{-}k} / 100)\} \times 100$
- SMM, Single Month Mortality
- Population Mean = Total sum/Total number of pools, Population SD = Population Standard Deviation

2010 and ‘Conforming loan’ since 2012 under the decline condition of mortgage rate in market is considered the mortgagor’s refinance incentive.

In table 4-2, the information about underlying assets of 205 FRM K-MBSs tells that the most housing type out of the assets is apartment (92.21%) and 60~70% of LTV (49.54%) is the largest level out of LTV distribution on the mortgagors. The underlying asset information on

Figure 4-1. Vintage analysis aging in month



time series base like payment histories based on LTV and geographic distribution, latest 12-month CPR history, and delinquency history cannot be gathered from KHFC DB, Bloomberg investor report, and Bloomberg terminal before 2014.

3. Variables

This study employed pool-level data that one K-MBS product is pooled with FRMs different from loan-level data. Compared with other MBS products like Fannie Mae having long history of operation, K-MBS applied in this study have short history, 13 years from 2004 to 2016. Moreover, accessing to loan-level data in KHFC is restrictive and useful data for analyzing prepayment factors are a few so that a model having simple structure and using small number of data is needed. Disclosed mortgage properties can be referred to KHFC DB based on Investor report. The disclosed data list is here,

- Underlying assets: Types, number of and amount of loans in each K-MBS pool

Table 4-2. Descriptive statistics of underlying assets on 205 K-MBS pools

	Number of houses	Population mean	Loan amount, KRW1million (Ratio %)	Population mean
Housing type				
Apartment	**1,572,639**	**7,671.4**	**156,630,788 (92.21)**	**764,052**
Multifamily housing	125,401	611.7	8,593,519 (5.06)	41,919
Row house	28,348	138.3	2,336,345 (1.38)	11,396
Single detached house	24,109	117.6	2,127,681 (1.25)	10,378
Others	1,802	8.8	176,291 (0.10)	859
Total	1,752,299	8,547.8	169,864,626 (100)	828,607
LTV distribution				
0-10%	6,426	31.3	170,269 (0.10)	830
10%-20%	56,982	278.0	2,273,603 (1.34)	11,090
20%-30%	119,855	584.7	6,788,429 (4.00)	33,114
30%-40%	168,975	824.3	12,370,488 (7.28)	60,343
40%-50%	212,438	1,036.3	18,926,269 (11.14)	92,323
50%-60%	440,278	2,147.7	45,181,119 (26.60)	220,395
60%-70%	**747,270**	**3,645.2**	**84,146,959 (49.54)**	**410,472**
70%-80%	51	0.2	5,818 (0.0034)	28
80%-90%	13	0.1	974 (0.0006)	4
90%-100%	3	0.0	225 (0.0001)	1
Total	1,752,299	8,548	169,864,626 (100)	828,607

Note.
- Below KRW 1million, Loan amount was cut
- Population mean = Total sum/Total number of pools

- Loan size: Issue amount of K-MBS, initial and current loan balance of underlying assets

- WARM, Weighted Average Residual Maturity and WA Seasoning

- WAC, Weighted Average Coupon

: The average of three five-year Treasury bonds from the past two days to current + Spread

- WAIR, Weighted Average Interest Rate

: Average interest rate of the underlying assets, weighted by current month's loan balance

- CPR, Conditional Prepayment Rate and SMM, Single Month Mortality

Based on the CLC document published since 2014, disclosed mortgage properties are here,

- The history of Originator, Geographic, Housing type, Original and Indexed LTV distributions

- Delinquency history

In order to reflect market environment factors related with K-MBS, this study obtained various market interest rate histories through 'ECOS (ECOnomic Statistics system)' of BOK[20] and KOFIA BIS (Bond Information Service)[21]. This study estimated the change rate of apartment purchase price as another key factor on K-MBS and acquired data about it from R-ONE (Korean real estate statistics) of KAB (Korea Appraisal Board)[22] and HPI (Housing Price Index) time series in month from KB (Kookmin Bank)[23].

This research assumes that the five main causes of the risk are refinance incentive caused by the change of interest rate, cash-out refinance, housing turnover, curtailment, and default that the K-MBS backed by FRMs has very low default risk. The prepayment risk influences the value of MBS so that accurate valuation of these factors and estimation of the risk amount can be the basis for promoting the mortgage market. In this study, the latest 12-month CPR is applied for prepayment as a dependent variable. The most important criterion in selecting an empirical model is how well the model can explain the relationship of the data. If a model is too complicated and takes too much time and cost to maintain, the use of the model will be limited.

For these reasons, this study adopted a parsimonious model, a pool-level linear model, based on the Schorin's model (1992) and made independent variables followed by two categories, refinance factors and the others. The market spread between WAIR of KHFC and the mortgage rate in housing market was used for estimating refinance incentive. For estimating cash-out

[20] BOK's ECOS
http://ecos.bok.or.kr/flex/visualStat2015.jsp
[21] KOFIA's BIS
http://www.kofiabond.or.kr/index.html
[22] KAB's R-ONE
http://www.r-one.co.kr/rone/resis/statistics/statisticsViewer.do,
[23] KB's HPI time series in month
http://nland.kbstar.com/quics?page=B046962

refinance, the logarithm changing rate of apartment purchase price is applied. In Figure 4-2, the annual averages of CPR, market spread (%), and the change rate of apartment purchase price are compared during the observation period from June 2004 to April 2018. As the average of market spread increases (the mortgage rate decreases) and decreases (the mortgage rate increases) after 2007, the average of CPR showed the similar movement with the mortgage rate.

When the comparison is considered with the change rate of apartment purchase price, prepayment can be explained more accurately. That is, the increase of apartment purchase price has refinance incentive under the decrease term of market spread. Housing turnover behaviors under this condition can also make more prepayment than under usual market condition. It is generally caused by seasonal movement. Most studies used the seasonal dummy variable that it will be 1 if the dummy is specific months in spring, fall and otherwise 0. In other to know the relationship between housing turnover and prepayment, the monthly average of SMM and the change rate of apartment purchase price are compared in Figure 4-3. As the large increase of apartment purchase price happened in March, April, October, and November, the average of SMM increased in the same months. This study selected March, April, October, and November for the seasonal dummy variable. The summary of variables is in Table 4-3.

Table 4-3. Descriptive statistics of variables in the total pool

Variable name	Observation	Mean	SD	Min	Max
CPR_{it}	14573	0.2832	0.1702	0	0.7356
Market $spread_{it}$	14573	1.3217	1.1844	-1.660	4.6400
$Aging_{it}$	14573	46.7746	35.7781	1	166
Log ΔHPI of Apt_t	14573	0.0009	0.0013	-.0038	0.0164

Note.
- $CPR_{it} = 1 - \prod_{k=0}^{11}(1 - smm_{12-k} / 100)$, Latest 12-month CPR; smm, single month mortality
- Market $spread_{it}$ = KHFC $WAIR_{it}$ – Average of mortgage $rate_t$ (Bank of Korea); WAIR, Weighted Average Interest Rate
- $Aging_{it}$ = Monthly aging since issue date
- Log ΔHPI of Apt_t = log(Apartment HPI_t) – log(Apartment HPI_{t-1}), Apartment HPI is Apartment purchase price index (% change) relative to December. 2015 which has HPI of 100

Figure 4-2. The trend comparison of refinance variables

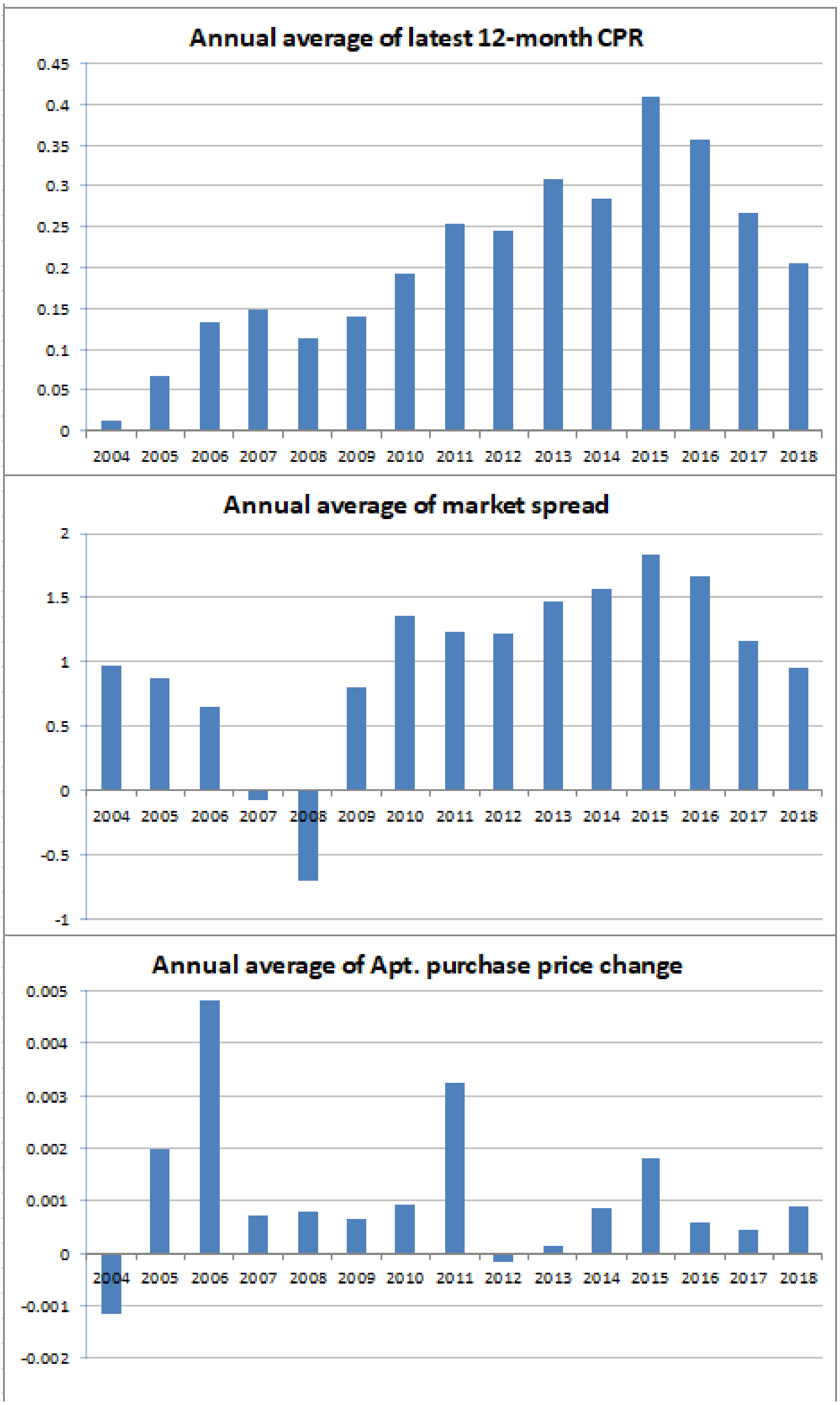
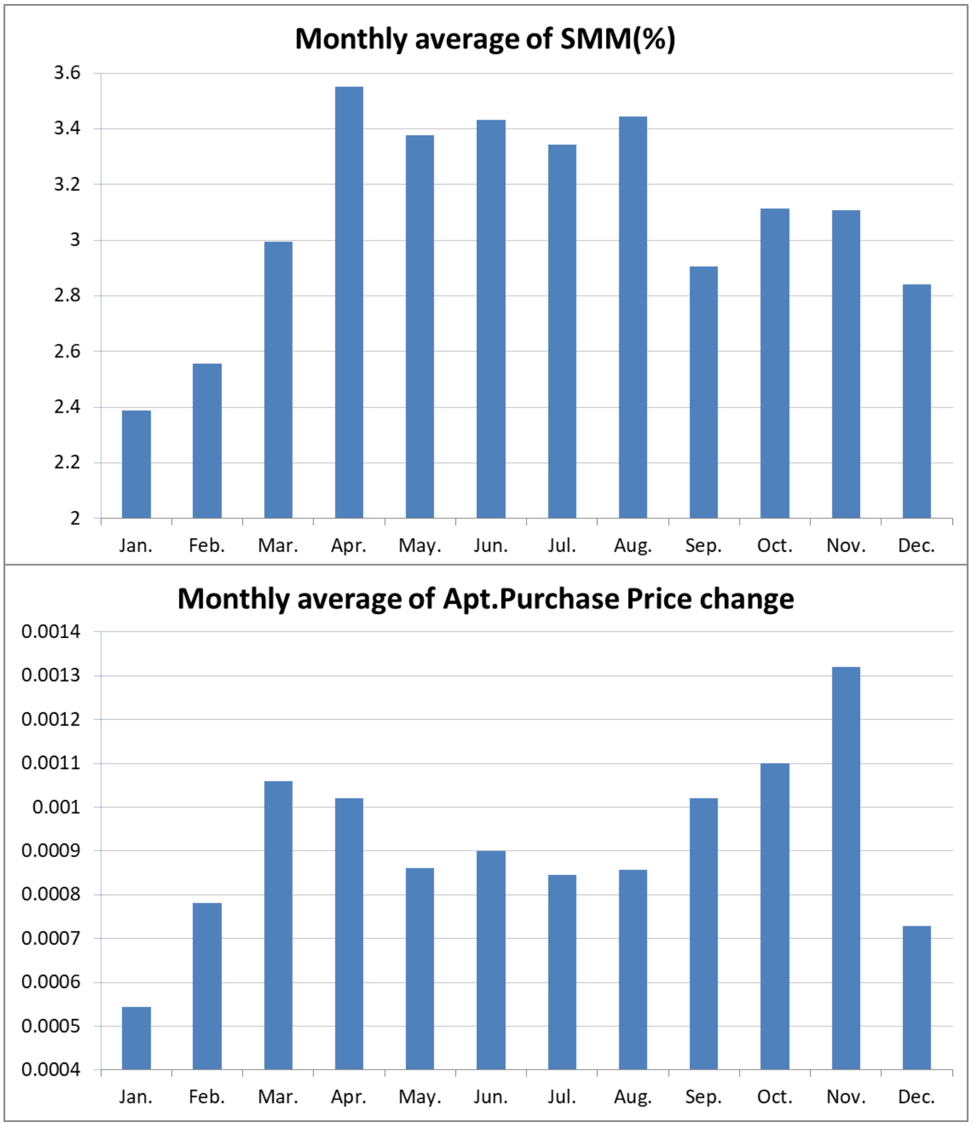


Figure 4-3. Movement season and Prepayment trend



4. Research model

As mentioned earlier, this study adopts a parsimonious model to directly estimate prepayment factors and independent variables are categorized into refinance factors and the other. The descriptive statistics of Variables are in Table 4-3. For making the panel data set, each annual pool based on each issue year from 2004 to 2016 of FRM K-MBSs was constructed at first. The total pool was created in the combination of these annual pools during the study observation from June 2004 to April 2018 and then regression analysis was performed with STATA. In spite of this effort to make significant variables on prepayment, this study also has unobserved factors from the data of KHFC DB. Next list of data has been disclosed since 2014. When KHFC cooperates in that kind of data before 2014, the study is possible to update the research model.

- The history of Originator, Geographic, Housing type, Original and Indexed LTV distributions
- Delinquency history

In the effort to control unobserved factors explained in ‘Chapter Ⅰ-4.Controlling unobserved factors’, this study applies panel data analysis with FE in addition to pooled OLS so that error term of OLS can be separated into two error terms, the time-invariant heterogeneity error term of K-MBS pools and the time-variant idiosyncratic error term of the pools. Moreover, ‘Within transformation’ of the FE is the ‘time-demeaning’ for each observation of each pool[19)]. The time-demeaning is equivalent to the both panel, balanced and unbalanced panel, in terms of obtaining the change in a dependent variable with the changes in independent variables by excluding the unobserved characteristics[24].

The most methodological disadvantage of pooled OLS is that it cannot control the unobserved heterogeneity. That is, researchers cannot recognize whether it is due to the independent variables or because of the idiosyncratic characteristics of the identity. Panel data analysis with FE based on ‘Within transformation’ can solve those problems like time-invariant heterogeneity such as the characteristics of loan characteristics (Bogeumjari, Didimdol,

[24] The critical issue with unbalanced panel is whether researchers know why it is unbalanced. Provided the reason why this study has unbalanced panel for pools is not correlated with the idiosyncratic error term, this study unbalanced panel can be used for consistent estimator.

Conforming, Ansim conversion loans for K-MBSs), data quality (documentation level), and difference level among K-MBS pools; the smaller number of loans the lower mortgagors.

Pooled OLS equation,

$Y_{i,t} = \alpha + \beta X_{i,t} + \varepsilon_{i,t}$; $\varepsilon_{i,t}$ is Error term, the unobserved heterogeneity

The brief form of panel regression (model),

$Y_{i,t} = \alpha + \beta X_{i,t} + u_i + e_{i,t}$

- u_i = Time-invariant heterogeneity error term of pools in panel

- $e_{i,t}$ = Time-variant idiosyncratic error term of pools in panel, pure error term

Between effects (model) in panel dataset,

$\bar{Y}_i = \alpha + \beta \dot{X}_i + u_i + \bar{e}_i$

- $\bar{Y}_i$, $\dot{X}_i$, $\bar{e}_i$ are the time-demeaning values or mean values of $Y_{i,t}$, $X_{i,t}$, $e_{i,t}$ of each identity

Using Within transformation, the difference between panel regression and Between effects,

$(Y_{i,t} - \bar{Y}_i) = \beta(X_{i,t} - \dot{X}_i) + (e_{i,t} - \bar{e}_i)$

- β is the expected change in the independent variable.

In elimination of u_i where time-invariant heterogeneity is included, consistent estimator of β can be obtained even in the condition $cov(X_{i,t}, u_i) \neq 0$

Another type of heterogeneity is called time heterogeneity. The study data set recorded in monthly time-series from the each issue date of each pool. This unbalanced panel is possible to have time heterogeneity caused by different time point. Even though the aging variable composed from each pool is applied for study models, 'Two-way FE' model with the time dummy variables, month dummy and year dummy, is considered. Normally, the equation with identity dummy can be written as

$Y_{i,t} = \sum_{i=1}^{n} (\alpha + u_i) + \beta X_{i,t} + e_{i,t}$

The two-way fixed effects model with time dummy can be written as

$Y_{i,t} = \sum_{i=1}^{n} (\alpha + u_i + v_t) + \beta X_{i,t} + e_{i,t}$; v_t is time heterogeneity

The models with dummy variables in this study have the occasions for interacting dummy variables with independent variables. The equation can be written as

$Y_{i,t} = \sum_{i=1}^{n} (\alpha + \delta_0 D_i) + (\beta + \delta_1 D_i) X_{i,t} + e_{i,t}$

D_i is dummy, δ_0 measures the difference in intercepts, and δ_1 measures the difference in explanatory variables.

Furthermore, in the case of K-MBSs pooled with single type of mortgages or otherwise, there are obviously some historical factors that make the dependent variable change in time series although it is difficult to explain the changes with other ways. For example, K-MBSs issued before and after GFC showed different prepayment speeds in Figure 4-1. Unobserved factors in this panel data set are possible to affect the both CPRs before and after GFC. In the study data set, Bogeumjari[25] loans have been provided to low and middle income households for relatively longer period than other loan types. Annual K-MBSs composed of Bogeumjari loans before

[25] t/e Bogeumjari loans have been securitized from 2004 to 2013 and u Bogeumjari loans have been securitized from 2010 to 2016. (Table 2-1)

GFC are more stable in cashflow than the others. Other K-MBSs pooled with several types of loans after GFC are prepaid faster than the other. It means that relatively older loans have been well known to mortgagors and younger loans vice versa. These unobserved historical factors have effect on the dependent variable such that it is auto-correlated within each panel group. This study employed one and two-month lagged[26] dependent variables as proxy independent variables for controlling the unobserved factors difficult to be explained in other ways.

Consequently, next three pooled OLS models as the base line models were constituted and performed at first. Three dummy variables (month dummy, year dummy, pool dummy) were added to each model in order to observe time and identity heterogeneities and variables are explained in Table 4-3. Since pooled OLS models cannot control the unobserved heterogeneity, the study doesn't consider $cov(CPR\_lag1_{i,t}$ or $CPR\_lag2_{i,t}, \varepsilon_{i,t}) \neq 0$; $\varepsilon_{i,t} = \rho\varepsilon_{i,t-1} + \eta_t$, the BLUE (Best Linear Unbiased Estimator) or consistent estimator assumption of OLS.

1) Pooled OLS No lagged CPR model

$$CPR_{i,t} = \alpha + \sum_k \beta_k \square X_{k\,i,t} + \varepsilon_{i,t}$$

2) Pooled OLS one-month lagged model or 1st order autoregressive (AR1) CPR model

$$CPR_{i,t} = \alpha + \sum_k \beta_k \square X_{k\,i,t} + \Gamma \square CPR\_lag1_{i,t} + \varepsilon_{i,t}$$

3) Pooled OLS two-month lagged model or 2nd order autoregressive (AR2) CPR model

$$CPR_{i,t} = \alpha + \sum_k \beta_k \square X_{k\,i,t} + \Gamma_1 \square CPR\_lag1_{i,t} + \Gamma_2 \square CPR\_lag2_{i,t} + \varepsilon_{i,t}$$

[26] AR2 CPR model has slightly different coefficients of constant and lagged terms from AR1 CPR model. Please refer to Wooldridge Jeffrey M. Introductory Econometrics - A Modern Approach, 5th edition, page 415-416.

The biggest disadvantage of pooled OLS is that it cannot control the unobserved heterogeneity so that next three FE models of panel data were constituted and performed. According to ‘Two-way FE’, two time dummy variables (month dummy, year dummy) were added to each model in order to observe time heterogeneity with time-invariant heterogeneity in the study panel data set. The study at first doesn’t consider ‘Dynamic panel regression models’ that uses one and two-month lagged dependent variables as proxy independent variables. It means $cov(CPR\_lag1_{i,t}$ or $CPR\_lag2_{i,t}, e_{i,t}) = 0$. Another correlation in the panel data, contemporaneous correlation, and the covariance matrix, homo- or heteroskedasticity of error term will be explained in Chapter Ⅳ-5.BLUE of ‘Dynamic panel regression models’.

1) FE No lagged CPR model

$$CPR_{i,t} = \alpha + \sum_k \beta_k \cdot X_{k\,i,t} + u_i + e_{i,t}$$

2) FE one-month lagged (AR1) CPR model

$$CPR_{i,t} = \alpha + \sum_k \beta_k \cdot X_{k\,i,t} + \Gamma \cdot CPR\_lag1_{i,t} + u_i + e_{i,t}$$

3) FE two-month lagged (AR2) CPR model

$$CPR_{i,t} = \alpha + \sum_k \beta_k \cdot X_{k\,i,t} + \Gamma_1 \cdot CPR\_lag1_{i,t} + \Gamma_2 \cdot CPR\_lag2_{i,t} + u_i + e_{i,t}$$

- u_i = Time-invariant heterogeneity error term of pools in panel
- e_{it} = Time-variant idiosyncratic error term of pools in panel, pure error term

Schorin’s analogy (1995) ‘pipeline lag’ (transportation lag) in system operation said that time differential among mortgagors obviously exists. Smarter mortgagors can understand the market situation and decide whether opening or closing a new mortgage and they influence successively the remaining mortgagors who are unaware of the situation. Therefore one and two month lag effects on the market spread (Market spread) and the change rate of apartment purchase price (Log ΔHPI of Apt.) are applied for three pooled OLS and three FE models.

According to the function of STATA (Statistics and data analysis tool) for panel data analysis with FE, 'xtreg Yvariable Xvariables, fe', the null hypothesis, $H_0 : u_i = 0$ for all i (annual MBS pool), was rejected from p-value of F-test. It means that BLUE cannot be obtained from pooled OLS. As a result, the nine FE models and the nine FE models with 2004~2018 year dummies[27] are employed for figuring out the inference analysis of the study. The nine FE models are here,

1) FE No lagged CPR model and the FE model added with one and two-month lagged independent variables.

2) FE one-month lagged (AR1) CPR model and the FE model added with one and two-month lagged independent variables.

3) FE two-month lagged (AR2) CPR model and the FE model added with one and two-month lagged independent variables.

Using STATA in this study can recognize whether these variables are jointly significant in p-value of F-statistic for the multi-collinearity that it might exist between the one and two-month lagged CPRs, between lagged CPRs and other explanatory variables, between the two main explanatory variables, the market spread and the logarithm changing rate of apartment purchase price, and their one and two-month lagged independent variables, and among explanatory variables.

[27] Since month dummies haven't shown significant effect, this study represents only year dummies for 'Two-way FE' model.

5. BLUE of ‘Dynamic panel regression models’

In econometrics sense, next four assumptions of BLUE are tackled in the study according to Gauss-Markov Theorem.

- Assumption1. $E(\varepsilon_{it}) = 0$, □ i and t
- Assumption2. $var(\varepsilon_{it}) = \sigma^2$, homoskedasticity, □ i and t
- Assumption3. $cov(\varepsilon_{it}, \varepsilon_{jt}) = 0$, □ i and t
- Assumption4. $cov(X_{it}, \varepsilon_{it}) = 0$, exogeneity of independent variables, □ i and t

Using panel data analysis with FE based on ‘Within transformation’, Assumption1 and 4 are verified in Chapter Ⅳ-4.Research model. In Assumption1, u_i, time-invariant heterogeneity error term ($\varepsilon_{it} = u_i + e_{it}$, time-variant idiosyncratic error term), is eliminated. Jointly significant p-value of F-statistic as a result of STATA allows the research models to satisfy Assumption4. In Assumption2, two heterogeneities, u_i (time-invariant heterogeneity) and v_t (time heterogeneity), are possible to coexist. It is controlled by Two-way Fixed Effects model with month dummy and year dummy variables.

Consequently, the study focuses on Assumption3. Since the study adds one and two-month lagged dependent variables to independent variables, the research models with these variables are considered ‘Dynamic panel regression models’. In Assumption3, there are two correlations of the error term, contemporaneous correlation and autocorrelation. In the case of contemporaneous correlation, it is included in u_i and STATA can treat it as heteroskedasticity with cross-sectional correlation in variance-covariance matrix.

When autocorrelation in ‘Dynamic panel regression models’ is suspicious, panel data analysis with FE based on ‘Within transformation’ cannot ensure BLUE for autocorrelation. To take one example of 1st order autoregressive (AR1) model,

$Y_{i,t} = \alpha + \gamma Y_{i,t-1} + \beta X_{i,t} + u_i + e_{i,t}$

In 'Within transformation', $cov(Y_{i,t-1} - \bar{Y}_i, e_{i,t} - \bar{e}_i) \neq 0$ since $e_{i,t-1}$ is included in $\bar{e}_i$. According to the function of STATA for panel data analysis with FE, 'xtreg Yvariable Xvariables, fe', this heterogeneity in panel data can be found in the proportional value, $\rho = 1/ (1 + \sigma_e^2 / \sigma_u^2)$. When ρ is close to 1 meaning that $\sigma_e^2 < \sigma_u^2$, control of u_i is significant in panel data.

For this reason, BLUE of 'Dynamic panel regression models' can be obtained by using additional method in panel data. IV (Instrument Variable) strategy that employs the past values of $Y_{i,t-1}$, $Y_{i,t-2}$, $Y_{i,t-3}$, ..., $Y_{i,t-T}$, is useful to control the correlation. However, FE-2SLS (Fixed Effects-2Stage Least squares) is not suitable for the strategy since $\bar{e}_i$ is composed of the past values of $e_{i,t}$, $e_{i,t-1}$, $e_{i,t-2}$, $e_{i,t-3}$, ..., $e_{i,t-T}$, and this IV strategy for FE-2SLS has the correlation problem between IVs and $\bar{e}_i$.

Generally, FD-2SLS (First Difference-2Stage Least squares) and Arellano-Bond estimations are applied for 'Dynamic panel regression models'. FD-2SLS and Arellano-Bond estimations have the first difference of error term, $\Delta e_{i,t} = e_{i,t} - e_{i,t-1}$, in process and IVs like $\Delta Y_{i,t-2}$, $\Delta Y_{i,t-3}$,

is relevant to the endogenous independent variable, $\Delta Y_{i,t-1}$ (instrument exogeneity) in the example of 1st order autoregressive (AR1) model.

In making use of IV, the next three conditions should be clear; 1) instrument relevance that IV is correlated with the endogenous independent variable; 2) instrument exogeneity that IV is uncorrelated with the error term in the original regression model; 3) IV is not included in the original regression model. This IV strategy for BLUE of 'Dynamic panel regression models' can be verified about the three conditions since the first difference of the endogenous variable, $\Delta Y_{i,t-1} = Y_{i,t-1} - Y_{i,t-2}$, is instrumented by the IV, $\Delta Y_{i,t-2} = Y_{i,t-2} - Y_{i,t-3}$. It makes FD-2SLS and

Arellano-Bond estimations verify these three conditions.

Especially, in the case of using ‘Dynamic panel regression models’ that uses one and two-month lagged dependent variables as proxy independent variables, the biggest advantage of Arellano-Bond estimation is that the process can generate itself the IVs for $\Delta Y_{i,t-1}$ and $\Delta Y_{i,t-2}$ unlike FD-2SLS. Arellano-Bond estimation based on system GMM (Generalized Method of Moments) is fitted to the over-identified model that the number of IVs are more than that of endogenous variables. Therefore the relevancy of Arellano-Bond estimation is tested by the serial correlation between the IVs and the error term instead of the over-identifying test as for testing the relevancy of the FD-2SLS. As a result, this study employs Arellano-Bond estimation in order to consider ‘Dynamic panel regression models’.

Ⅴ. Result

1. Non-dynamic panel regression models, cov($CPR\_lag1_{i,t}$ or $CPR\_lag2_{i,t}$, $e_{i,t}$) = 0

The result of nine FE models in the total pool of K-MBSs (Table 5-1) compares the FE No lagged CPR (model)s with FE AR1, AR2 (1^{st}, 2^{nd} order autoregressive) CPR (model)s. After the time lag effects from one month to 12 months on the dependent variable (CPR) and independent variables (Market spread and Log ΔHPI of Apt.) were analyzed, the most significant AR1, AR2 CPRs and one and two-month lagged independent variables are shown in Table 5-1. Overall, all explanatory variables except 'Season dummy' are significant. In the case of 'Aging' in month, it seems to be mostly significant to positive effect but the coefficient is too low. As a result of F-statistic for the multi-collinearity, all models are jointly significant.

'Market spreads', the most important explanatory variable, are significant in all nine FE models. Lagged Market Spreads are significant in all six models of one and two-month lagged independent variables. No lagged CPRs with one and two-month lagged Market spreads have similar effects between Market spread in current month and Lagged Market spread in one month, two months earlier while AR1 and AR2 CPRs with one and two-month lagged Market spreads have opposite effects.

'Log ΔHPIs of Apt.' are also significant in all nine FE models and Lagged Log ΔHPIs of Apt. are significant in three models except AR1 and AR2 CPRs in one-month lagged independent variables and AR1 CPR in two-month lagged independent variables. No lagged CPR in one-month lagged independent variables shows opposite effect between Log ΔHPI of Apt. in current month and Lagged Log ΔHPI of Apt. in one month earlier. No lagged CPR and AR2 CPR in two-month lagged independent variables show also opposite effects between Log ΔHPI of Apt. in current month and Lagged Log ΔHPI of Apt. in two months earlier.

Table 5-1. Coefficients on nine FE models of panel data (14,573 total observations)

Lagged independent variables	NO lagged independent variables			One-month lagged independent variables			Two-month lagged independent variables		
9 FE models	No lagged CPR	AR1 CPR_{it-1}	**AR2 CPR_{it-2}**	No lagged CPR	AR1 CPR_{it-1}	AR2 CPR_{it-2}	No lagged CPR	AR1 CPR_{it-1}	AR2 CPR_{it-2}
Market $spread_{it}$	0.0951 *** (0.0021)	0.0098 *** (0.0002)	**0.0057 *** (0.0002)**	0.0423 *** (0.0089)	0.0188 *** (0.0011)	0.0173 *** (0.0009)	0.0569 *** (0.0053)	0.0169 *** (0.0007)	0.0121 *** (0.0006)
Lagged Market $spread_{it}$				0.0590 *** (0.0091)	-0.0094 *** (0.0012)	-0.0124 *** (0.0010)	0.0502 *** (0.0055)	-0.0079 *** (0.0007)	-0.0076 *** (0.0006)
Log ΔHPI of Apt_t	6.39863 *** (0.7799)	1.06586 *** (0.1030)	**0.44339 *** (0.0867)**	-6.46470 *** (1.3331)	1.04975 *** (0.1764)	0.91240 *** (0.1471)	-4.22509 *** (0.9611)	1.44907 *** (0.1294)	0.99431 *** (0.1069)
Lagged Log ΔHPI of Apt_t				13.1327 *** (1.3177)	0.3209 (0.1748)	-0.1712 (0.1458)	12.3472 *** (0.9497)	-0.0415 (0.1285)	-0.3288 ** (0.1060)
$Aging_{it}$	0.0004 *** (0.00006)	-0.0003 *** (0.00001)	**-0.0001 *** (0.00001)**	0.0002 ** (0.00006)	-0.0003 *** (0.00001)	-0.0001 *** (0.00001)	-0.0000 (0.00007)	-0.0003 *** (0.00001)	-0.0001 *** (0.00001)
Season $dummy_i$	-0.0053 ** (0.0020)	0.00006 (0.0002)	**0.0001 (0.0002)**	-0.0017 (0.0020)	-0.00006 (0.0002)	-0.00009 (0.0002)	0.0018 (0.0020)	-0.0002 (0.0002)	-0.0002 (0.0002)
AR1 CPR_{it-1} of CPR		0.9805 *** (0.0011)	**1.5375 *** (0.0070)**		0.9810 *** (0.0011)	1.5429 *** (0.0070)		0.9807 *** (0.0011)	1.5400 *** (0.0069)
AR2 CPR_{it-2} of CPR			**-0.5579 *** (0.0069)**			-0.5622 *** (0.0069)			-0.5586 *** (0.0069)
σ_u	0.0529	0.0043	**0.0021**	0.0526	0.0041	0.0018	0.0526	0.0039	0.0016
σ_e	0.1172	0.0153	**0.0127**	0.1155	0.0152	0.0126	0.1138	0.0153	0.0126
ρ	0.1695	0.0733	**0.0277**	0.1720	0.0678	0.0209	0.1759	0.0638	0.0174
Adjusted R^2	0.3784	0.9890	**0.9921**	0.3760	0.9890	0.9922	0.3739	0.9886	0.9922
AIC	-21326.7	-79511.4	**-83639.7**	-21430.2	-79570.1	-83801.0	-21547.8	-78380.2	-83825.7
F-test, H_0: $u_i = 0$						H_0			H_0

Note.
- Standard errors are in parentheses, * p<0.05, ** p<0.01, *** p<0.001
- $CPR_{it} = 1 - \prod_{k=0}^{11}(1 - smm_{12-k} / 100)$, Latest 12-month CPR; smm, single month mortality
- Market $spread_{it}$ = KHFC $WAIR_{it}$ – Average of mortgage $rate_t$ (Bank of Korea); WAIR, Weighted Average Interest Rate
- $Aging_{it}$ = Monthly aging since issue date
- Log ΔHPI of Apt_t = log(Apartment HPI_t) – log(Apartment HPI_{t-1}), Apartment HPI is Apartment purchase price index (% change) relative to December. 2015 which has HPI of 100
- Season $dummy_i$ = Movement season dummy. If it is March, April, October and November the value will be 1 and otherwise, the value is 0
- $\rho = 1/ (1 + \sigma_e^2 / \sigma_u^2)$. When ρ is close to 1 meaning that $\sigma_e^2 < \sigma_u^2$, control of u_i is significant in panel data
- AR1, AR2 of CPR_{it} = 1, 2-month lagged autoregressive variables of CPR
- AIC = Akaike Information Criterion
- F-test, H_0: $u_i = 0$, is to test model fitness between pooled OLS (H_0) and Fixed Effects in panel data

Because of the weakness of 'Within transformation', $cov(Y_{i,t-1} - \bar{Y}_i, e_{i,t} - \bar{e}_i) \neq 0$ since $e_{i,t-1}$ is included in $\bar{e}_i$, this unobserved heterogeneity in panel data can be found in the proportional value, $\rho = 1/ (1 + \sigma_e^2 / \sigma_u^2)$. In Table 5-1 and 5-2, almost all FE models show that σ_e and σ_u, are quite small so that the study analyzes at first non-dynamic panel regression models, $cov(CPR\_lag1_{i,t}$ or $CPR\_lag2_{i,t}, e_{i,t}) = 0$.

The study also considers this unobserved heterogeneity of 'Dynamic panel regression models' with Arellano-Bond estimation. According to adjusted R^2, AIC[28] that is an estimator for the relative quality of statistical models in a given data set, and F-test for H_0: $u_i = 0$, Arellano-Bond estimation is applied to the optimal model, the AR2 CPR model in NO lagged independent variables.

Based on adjusted R^2, AIC, and F-test of the result, the explanatory levels of No lagged CPRs are not significantly better than those of AR1 and AR2 CPRs. Comparing the coefficients of AR1, AR2 CPRs in NO lagged independent variables with those in one and two-month lagged independent variables, the lag effects on CPRs are similar in all three models of AR1 CPR and three models of AR2 CPR regardless of F-test for H_0: $u_i = 0$.

The result of adjusted R^2, AIC, and F-test shows that the AR2 CPR model in NO lagged independent variables is the most significant model that explains next the effects of independent variables.

1) When Market spread increases 1% point, CPR increases 0.57% point.

2) In the case of Log ΔHPI of Apt., the interpretation is derived from level-log model,

$y = \alpha + \beta \log x$

---

[28] AIC, Akaike Information Criterion = 2k – 2Ln(L). The preferred model has the minimum value of AIC.
- k is the number of estimated parameters in the model
- L is the maximum value of the likelihood function for the model

$y + \Delta y = \alpha + \beta \log(x + \Delta x)$; y will be subtracted

$\Delta y = \beta\{\log(x + \Delta x) - \log x\} = \beta \log(1 + \Delta x/x)$

If $\Delta x/x$ is small, we can apply $\log(1 + \Delta x/x) \approx \Delta x/x$

$\Delta y \approx \beta(\Delta x/x)$ or $\Delta y \approx \beta/100(100*\Delta x/x)$

It means that 1% change of HPI of Apt_t from the previous HPI of Apt_{t-1} affects β/100 change of CPR. When ΔHPI of Apt. increases 1% point, CPR increases 0.44339% point.

3) AR2 CPR model implies that 1% point increase in CPR one month ago affects 1.5375% point increase in the CPR this year and that of CPR two months ago lead to an estimated 0.5579% point drop in the CPR this year.

2004~2018 year dummies are applied to the previous nine FE models in the total pool of FRM K-MBSs and the results are shown in Table 5-2. Overall, all explanatory variables except 'Season dummy' are significant. In the case of 'Aging' in month, it seems to be mostly significant to positive effect but the coefficient is too low. As a result of F-statistic for the multi-collinearity, all models are jointly significant.

'Market spreads', the most important explanatory variable, are significant in eight models except No lagged CPR in one-month lagged independent variables. Lagged Market Spreads are significant in four models except AR1 CPRs in one and two-month lagged independent variables. AR2 CPR in one-month lagged independent variables has opposite effect between Market spread in current month and Lagged Market spread in one month earlier. In two-month lagged independent variables, No lagged CPR has similar effect between Market spread in current month and Lagged Market spread in two months earlier while AR2 CPR has opposite effect.

‘Log ΔHPIs of Apt.’ are significant in seven models except No lagged CPRs in one and two-month lagged independent variables. Lagged Log ΔHPIs of Apt. are significant in four models except AR1 and AR2 CPRs in one-month lagged independent variables. AR1 and AR2 CPRs in two-month lagged independent variables show opposite effects between Log ΔHPI of Apt. in current month and Lagged Log ΔHPI of Apt. in two months earlier.

Based on adjusted R^2, AIC, and F-test of the result, the explanatory levels of No lagged CPRs are not significantly better than those of AR1 and AR2 CPRs. Comparing the coefficients of AR1, AR2 CPR models in NO lagged independent variables with those in one and two-month lagged independent variables, the lag effects on CPRs are similar in all three models of AR1 CPR and three models of AR2 CPR regardless of F-test for H_0: $u_i = 0$.

The result of adjusted R^2, AIC, and F-test shows that the AR2 CPR model in NO lagged independent variables is the most significant model that explains next the effects of independent variables.

1) When Market spread increases 1% point, CPR increases 1.07% point.

2) When ΔHPI of Apt. increases 1% point, CPR increases 0.6405% point.

3) AR2 CPR model implies that 1% point increase in CPR one month ago affects 1.4519% point increase in the CPR this year and that of CPR two months ago lead to an estimated 0.4732% point drop in the CPR this year.

Table 5-2. Coefficients on nine FE models of panel data with 2004~2018 year dummies

Lagged independent variables	NO lagged independent variables			One-month lagged independent variables			Two-month lagged independent variables		
9 FE models	No lagged CPR	AR1 CPR_{it-1}	**AR2 CPR_{it-2}**	No lagged CPR	AR1 CPR_{it-1}	AR2 CPR_{it-2}	No lagged CPR	AR1 CPR_{it-1}	AR2 CPR_{it-2}
Market $spread_{it}$	0.0245 *** (0.0050)	0.0143 *** (0.0006)	**0.0107** *** **(0.0005)**	0.0023 (0.0094)	0.0133 *** (0.0012)	0.0164 *** (0.0010)	0.0192 ** (0.0064)	0.0144 *** (0.0008)	0.0125 *** (0.0007)
Lagged Market $spread_{it}$				0.0311 *** (0.0091)	0.0012 (0.0011)	-0.0069 *** (0.0010)	0.0214 *** (0.0058)	0.0001 (0.0007)	-0.0036 *** (0.0006)
Log ΔHPI of Apt_t	2.3320 * (0.9996)	1.0757 *** (0.1291)	**0.6405** *** **(0.1147)**	-2.1728 (1.3467)	0.9750 *** (0.1744)	0.9938 *** (0.1541)	-0.9588 (1.0804)	1.2452 *** (0.1417)	1.0281 *** (0.1243)
Lagged Log ΔHPI of Apt_t				5.3356 *** (1.3186)	0.0944 (0.1708)	-0.2226 (0.1513)	6.0876 *** (1.0025)	-0.4444 *** (0.1317)	-0.5699 *** (0.1155)
$Aging_{it}$	0.0004 (0.0002)	-0.0004 *** (0.00004)	**-0.0002** *** **(0.00003)**	-0.0001 (0.00029)	-0.0004 *** (0.00004)	-0.0002 *** (0.00003)	-0.000 (0.00030)	-0.0004 *** (0.00004)	-0.0002 *** (0.00003)
Season $dummy_i$	-0.0012 (0.0019)	-0.0000 (0.0002)	**0.0001** **(0.0002)**	0.0003 (0.0019)	0.00004 (0.0002)	-0.00007 (0.0002)	0.0015 (0.0019)	-0.0001 (0.0002)	-0.0002 (0.0002)
AR1 CPR_{it-1} of CPR		0.9792 *** (0.0010)	**1.4519** *** **(0.0074)**		0.9791 *** (0.0010)	1.4590 *** (0.0075)		0.9776 *** (0.0011)	1.4576 *** (0.0074)
AR2 CPR_{it-2} of CPR			**-0.4732** *** **(0.0073)**			-0.4800 *** (0.0074)			-0.4783 *** (0.0073)
σ_u	0.0982	0.0083	**0.0059**	0.1033	0.0083	0.0064	0.0930	0.0084	0.0055
σ_e	0.1093	0.0140	**0.0123**	0.1082	0.0140	0.0123	0.1069	0.0140	0.0123
ρ	0.4466	0.2605	**0.1910**	0.4771	0.2600	0.2148	0.4306	0.2651	0.1698
Adjusted R^2	0.4590	0.9908	**0.9926**	0.4531	0.9908	0.9926	0.4478	0.9905	0.9927
AIC	-23336.5	-82053.4	**-84504.1**	-23311.6	-82051.0	-84549.7	-23312.1	-80841.8	-84569.1
F-test, H_0: $u_i = 0$						H_0 at *			H_0

Note.

- Standard errors are in parentheses, * $p<0.05$, ** $p<0.01$, *** $p<0.001$
- $CPR_{it} = 1 - \prod_{k=0}^{11}(1 - smm_{12-k} / 100)$, Latest 12-month CPR; smm, single month mortality
- Market $spread_{it}$ = KHFC $WAIR_{it}$ – Average of mortgage $rate_t$ (Bank of Korea); WAIR, Weighted Average Interest Rate
- $Aging_{it}$ = Monthly aging since issue date
- Log ΔHPI of Apt_t = log(Apartment HPI_t) – log(Apartment HPI_{t-1}), Apartment HPI is Apartment purchase price index (% change) relative to December. 2015 which has HPI of 100
- Season $dummy_i$ = Movement season dummy. If it is March, April, October and November the value will be 1 and otherwise, the value is 0
- $\rho = 1/ (1 + \sigma_e^2 / \sigma_u^2)$. When ρ is close to 1 meaning that $\sigma_e^2 < \sigma_u^2$, control of u_i is significant in panel data
- AR1, AR2 of CPR_{it} = 1, 2-month lagged autoregressive variables of CPR
- AIC = Akaike Information Criterion
- F-test, H_0: $u_i = 0$, is to test model fitness between pooled OLS (H_0) and Fixed Effects in panel data

2. Dynamic panel regression models, cov(CPR_lag1$_{i,t}$ or CPR_lag2$_{i,t}$, $e_{i,t}$) ≠ 0

In addition to the fact that almost all FE models show that σ_e and σ_u, are quite small, the study also considers this unobserved heterogeneity of 'Dynamic panel regression models' with Arellano-Bond estimation. In STATA function considering heteroskedasticity, 'xtabond Yvariable Xvariables, lag(2) nocons robust', BLUE assumptions according to Gauss-Markov Theorem are verified.

According to adjusted R^2, AIC, and F-test, Arellano-Bond estimation is applied to the optimal model, the AR2 CPR model in NO lagged independent variables. Comparing with the result of the AR2 CPR model in NO lagged independent variables (Table 5-1), one and two-month lagged CPR show the similar effects so that the study assumption of non-dynamic panel regression models, cov(CPR_lag1$_{i,t}$ or CPR_lag2$_{i,t}$, $e_{i,t}$) = 0, is reasonable.

Figure 5-1. The result of Arellano-Bond estimation

```
Arellano-Bond dynamic panel-data estimation     Number of obs      =     13,958
Group variable: id                              Number of groups   =        205
Time variable: t
                                                Obs per group:
                                                              min =         14
                                                              avg =    68.0878
                                                              max =        163

Number of instruments =    5.9e+03              Wald chi2(6)       =   1.07e+06
                                                Prob > chi2        =     0.0000
One-step results
------------------------------------------------------------------------------
         cpr |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         cpr |
         L1. |   1.502401   .0075627   198.66   0.000     1.487578    1.517224
         L2. |  -.5259171   .0074285   -70.80   0.000    -.5404767   -.5113576
             |
market_spread |   .0070291   .0003327    21.13   0.000      .006377    .0076812
       aging |  -.0002377   9.07e-06   -26.21   0.000    -.0002555   -.0002199
logdeltahpi_apt |   .6543369   .1150309     5.69   0.000     .4288805    .8797934
  dum_season |   .0002324   .0002401     0.97   0.333    -.0002381     .000703
------------------------------------------------------------------------------
Instruments for differenced equation
        GMM-type: L(2/.).cpr
        Standard: D.market_spread D.aging D.logdeltahpi_apt D.dum_season
```

Ⅵ. Policy implication

The MBS market plays an important role in the bond market around the world and the MBS market is closely related to the housing finance market since MBS securitizes mortgages as underlying assets. Therefore, in order to build an efficient MBS market and activate the housing finance market, the basic research on housing finance products like this study is essential and improvement schemes and programs should be established to make MBS studies systematic and sustainable.

The role of developing MBS market in Korea depends largely on KHFC. As mentioned earlier however, micro information DB of K-MBS including housing information is not completely established so that more specific analyses linked with the housing market is limited. For this reason, the study suggests some policy implications focused on enhancing MBS market in Korea.

1) For monitoring housing finance market systematically and establishing integrated risk management system, KHFC needs to build housing finance information system.

□ Construction of micro data through the combination among characteristics of mortgages, houses, and mortgagors

□ Based on the KHFC's research DB that the linkage analysis between the housing market and the financial environment will be provided, KHFC can strengthen the monitoring and policy support function of the housing finance market

□ Sharing data and information with associated agencies for eliminating information asymmetry among government, financial institutions, and financial consumers can make the housing finance market transparent and stable

2) Upgrading MBS transaction information and infrastructure.

□ The risk management infrastructure such as prepayment-OAS and credit risk indicators for the optional MBS with five-year maturity or more is need

□Updating CLTV (Current Lone to Value) and CDTI (Current Debts to Income) periodically in connection with the price information of underlying assets and the income information of mortgagors

□ Considering expansion of various investors (foreign investors included) interested in securities backed by RP (Repurchase agreement) bond of BOK or high quality liquid assets (Lowest level at liquidity coverage ratio)

□ Raising MBS primary dealers

3) Policy effect analysis of policy mortgages.

□ Analyzing the mortgagors' behavior after loan such as additional loans (fixed or floating rate), card consumption, and income trend

□ Identifying whether a second home has been purchased or just movement after loan through actual price matching

□ Studying sample groups that have similar characteristics to the FRM mortgagors and analyzing key factors on the selection of policy mortgages or general mortgages

Ⅶ. Conclusion

This study was conducted to analyze the factors of prepayment risk in K-MBSs issued by KHFC. Total 205 FRM K-MBSs issued from 2004 to 2016 and securitized with total 1,752,299 underlying assets are used and 14,573 observations were recorded monthly from June 2004 to April 2018 by KHFC. Four independent variables (the market spread, the change rate of apartment purchase price, aging in month, and seasonal movement dummy) are estimated for the effects on the dependent variable, latest 12-month CPR.

One of pool-level linear models based on the Schorin's method (1992, 1995) is chosen as a parsimonious model for direct estimation of prepayment factors. In order to control several unobserved factors, panel data analysis with FE is performed and it is added with 2004~2018 year dummies, with one and two-month lag effects on the market spread and the logarithm changing rate of apartment purchase price, and with one and two-month lag effects of CPR.

Overall, all explanatory variables except 'Season' dummy are significant. In the case of 'Aging' in month, it seems to be mostly significant but the coefficient is too low. Based on the result of adjusted R^2, AIC, and F-test for H_0: $u_i = 0$, the most significant models in both nine FE models and nine FE models with 2004~2018 year dummies are the AR2 CPR models in NO lagged independent variables. The result summary of the AR2 CPR models is here,

1) When the market spread increases 1% point, CPR increases 0.57 ~ 1.07% point. Holding other factors constant, the hypothesis that prepayment (CPR) will increase when the market spread is expanded is proved.

2) When ΔHPI of Apt. increases 1% point, CPR increases 0.4433 ~ 0.6405% point. It means that the price increase caused by cash-out refinance or housing turnover behavior can make prepayment increase when holding other patterns constant.

3) AR2 CPR model implies that 1% point increase in CPR one month ago affects 1.4519 ~ 1.5375% point increase in the CPR this year and that of CPR two months ago lead to an estimated 0.4732 ~ 0.5579% point drop in the CPR this year.

In both nine FE models and nine FE models with 2004~2018 year dummies, all independent variables are jointly significant in p-value of F-statistic and the time lag effects on CPR are similar in all six AR1 CPR models and all six AR2 CPR models regardless of F-test for H_0: $u_i = 0$ so that the time lag effects on the dependent variable (CPR) exist in this research whether year dummies are applied to the models or not.

Comparing the coefficients of AR1, AR2 CPRs in NO lagged independent variables with those in one and two-month lagged independent variables, it is also considered that there might be dispersion of the time lag effects on the dependent variable, the market spread, and the logarithm changing rate of apartment purchase price even though the models are jointly significant in p-value of F-statistic.

Arellano-Bond estimation is applied to the optimal model, the AR2 CPR model in NO lagged independent variables. Comparing with the result of the AR2 CPR model in NO lagged independent variables (Table 5-1), one and two-month lagged CPR show the similar effects so that the study assumption of non-dynamic panel regression models, cov($CPR\_lag1_{i,t}$ or $CPR\_lag2_{i,t}$, $e_{i,t}$) = 0, is reasonable.

Consequently, the pool-level of this study with some controlling measures about the unobserved factors in the panel dataset has an advantage that the study method estimates easily and directly the main four variables on CPR compared with other method introduced in the literature review. This study will enhance the understanding of the prepayment factors of K-MBS and will be the basis for pricing K-MBS.

Reference

1) Schorin, C.N.(1992), Modeling and projecting MBS prepayments in Handbook of Mortgage Backed Securities, edited by Frank J. Fabozzi.

2) Schorin, C.N.(1995), Mortgage prepayment modeling 1 in Handbook of Mortgage Backed Securities, edited by Frank J. Fabozzi.

3) Korea Housing Institute (June 2015), Present state and issues of low and mid-income household housing finance.

4) Man Cho(2012), Housing finance, macro- and microeconomic prudence regulation: Institutional and empirical analysis for developing market indicators and policies in New paradigm of housing policy in Korea housing market (KDI research report series).

5) Korea Securities Research Institute (July 2008), A study of the KHFC MBS valuation based on prepayment analysis: OAS Approach.

6) Asay, M., F. H., Guillaume and R. K., Mattu (1987), Duration and Convexity of Mortgage Backed Securities: Some Hedging Implications from a Prepayment Linked Present Value Model, in Mortgage Backed Securities, edited by Frank J. Fabozzi.

7) Brazil, A. J. (1988), Citicorp's Mortgage Valuation Model: Option-Adjusted Spreads and Option Based Duration, Journal of Real Estate and Economics.

8) Carron, A. S. and M. Hogan (1988), The Option Valuation Approach to Mortgage Pricing, Journal of Real Estate Finance and Economics.

9) Davidson, A., M. Herskovits and L. D. Van Drunen (1988), The Refinancing Threshold Pricing Model: An Economic Approach to Valuing MBS, Journal of Real Estate Finance and Economics.

10) Chinloy, P. (1991), The Option Structure of a Mortgage Contract, Journal of Housing Research.

11) Richard, S. F. and R. Roll (1989), Prepayments on Fixed-rate Mortgage-Backed Securities, The Journal of Portfolio Management.

12) OFHEO Working paper 02-1 (March 2002), Patterns of Default and Prepayment for Prime and Nonprime Mortgages.

13) Schwartz, E. S. and W. N. Torous(1989), Prepayment and the Valuation of Mortgage-Backed Securities, Journal of Finance.

14) Cox, D. R. (1972), Regression Models and Life-Tables, Journal of the Royal Statistical Society Series.

15) Cox, D. R. (1975), Partial Likelihood, Biometrica.

16) Hosmer, David W (2008), Applied survival analysis: Regression modeling of time-to-event data, 2nd edition.

17) Cleves, Mario A. (2004), An Introduction to survival analysis using Stata, Revised edition.

18) Housing Finance Research Institute of KHFC (December 2017), Analysis and utilization of K-MBS DB, Housing finance, policy report.

19) Wooldridge Jeffrey M. Introductory Econometrics: A Modern Approach, 5th edition, page 491.

Appendix

Table A. Data description of t/e and u Bogeumjari

Loan amount (KRW 1million)	t/e Bogeumjari (2004-2013)		u Bogeumjari (2010-2013)	
	Total	Population mean	Total	Population mean
	23,216,733	386,945	26,668,519	555,594
Total current LB (KRW 1million), 2018.04	247,271	4,121	2,812,224	58,588
WARM (month)		119.7		127.0
Number of loans	284,384	4,739.7	252,187	5,253.9
Number of pools	60	6 (Annual mean)	48	12 (Annual mean)
Initial WAIR (%)		6.22		4.51
Current WAIR (%), 2018.04		5.72		4.43
WAC (%)		5.08		3.62
Origination spread (%)		1.13		0.89
Avg. CPR		0.32		0.31
Avg. SMM (%)		3.45		3.36

Note.
- Below KRW 1million, Loan amount was cut
- Population mean = Total sum/Total number of pools

Table B. Asset information of t/e and u Bogeumjari (Continued)

Housing type	t/e Bogeumjari (2004-2013), 60 pools		u Bogeumjari (2010-2013), 48 pools	
Number of houses	Total	Population mean	Total	Population mean
Apartment	**276,204**	**4,603.4**	**244,096**	**5,085.3**
Multifamily Housing	4,208	70.1	4,837	100.8
Row House	2,169	36.2	1,698	35.4
Single Detached House	1,803	30.1	1,400	29.2
Others	0	0	156	3.3
Total	284,384	4,739.7	252,187	5,253.9
Loan amount (KRW 1million)	Total (Ratio %)	Population mean	Total (Ratio %)	Population mean
Apartment	**22,708,298 (97.81)**	**378,471**	**26,190,525(98.21)**	**545,635**
Multifamily Housing	226,444 (0.98)	3,774	255,267 (0.96)	5,318
Row House	158,083 (0.68)	2,634	108,791 (0.41)	2,266
Single Detached House	123,907 (0.53)	2,065	97,401 (0.37)	2,029
Others	0	0	16,533 (0.06)	344
Total	23,216,733 (100)	386,945	26,668,519 (100)	555,594

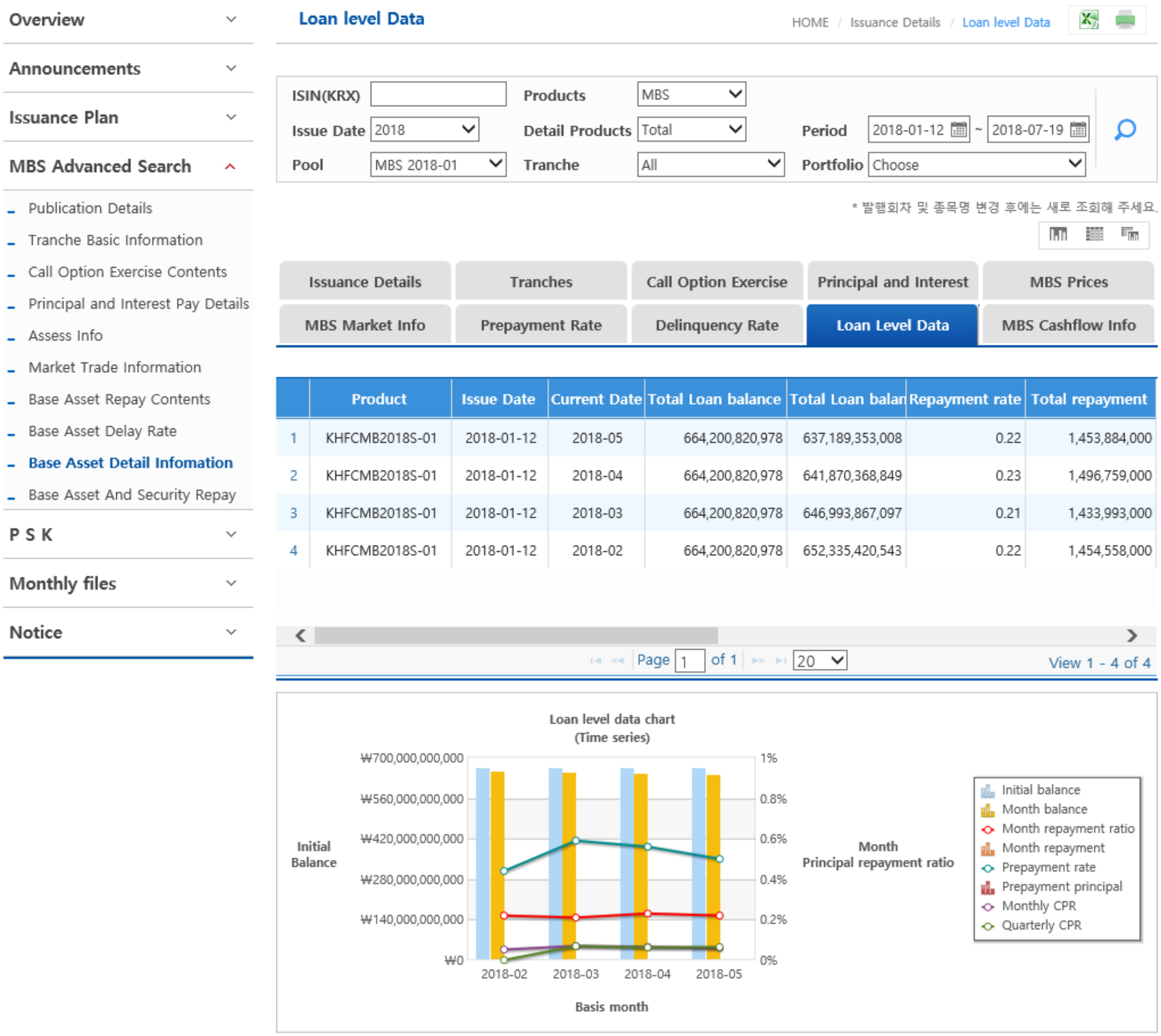
Note.
- Below KRW 1million, Loan amount was cut
- Population mean = Total sum/Total number of pools

Table B. Asset information of t/e and u Bogeumjari

LTV Distribution	t/e Bogeumjari (2004-2013), 60 pools		u Bogeumjari (2010-2013), 48 pools	
Number of houses	Total	Population mean	Total	Population mean
0 -10%	1,130	18.8	819	17.1
10%-20%	6,033	100.6	6,361	132.5
20%-30%	11,984	199.7	13,274	276.5
30%-40%	18,492	308.2	20,627	429.7
40%-50%	28,274	471.2	26,957	561.6
50%-60%	88,616	1,476.9	61,773	1,286.9
60%-70%	**129,855**	**2,164.3**	**122,375**	**2,549.5**
Total	284,384	4,739.7	252,187	5,253.9
Loan amount (KRW 1million)	Total (Ratio %)	Population mean	Total (Ratio %)	Population mean
0-10%	43,349 (0.19)	722	17,525 (0.07)	365
10%-20%	238,119 (1.03)	3,968	240,576 (0.90)	5,012
20%-30%	665,554 (2.87)	11,092	735,562 (2.76)	15,324
30%-40%	1,266,147 (5.45)	21,102	1,487,723 (5.58)	30,994
40%-50%	2,176,179 (9.37)	36,269	2,318,808 (8.69)	48,308
50%-60%	7,928,131 (34.15)	132,135	6,707,506 (25.15)	139,739
60%-70%	**10,899,251 (46.95)**	**181,654**	**15,160,715 (56.85)**	**315,848**
Total	23,216,733(100)	386,945	26,668,519 (100)	555,594

Note.
- Below KRW 1million, Loan amount was cut
- Population mean = Total sum/Total number of pools

Figure A. Loan Level Data in KHFC DB



	Product	Issue Date	Current Date	Total Loan balance	Total Loan balan	Repayment rate	Total repayment
1	KHFCMB2018S-01	2018-01-12	2018-05	664,200,820,978	637,189,353,008	0.22	1,453,884,000
2	KHFCMB2018S-01	2018-01-12	2018-04	664,200,820,978	641,870,368,849	0.23	1,496,759,000
3	KHFCMB2018S-01	2018-01-12	2018-03	664,200,820,978	646,993,867,097	0.21	1,433,993,000
4	KHFCMB2018S-01	2018-01-12	2018-02	664,200,820,978	652,335,420,543	0.22	1,454,558,000

Figure B. MBS Cashflow Information in KHFC DB

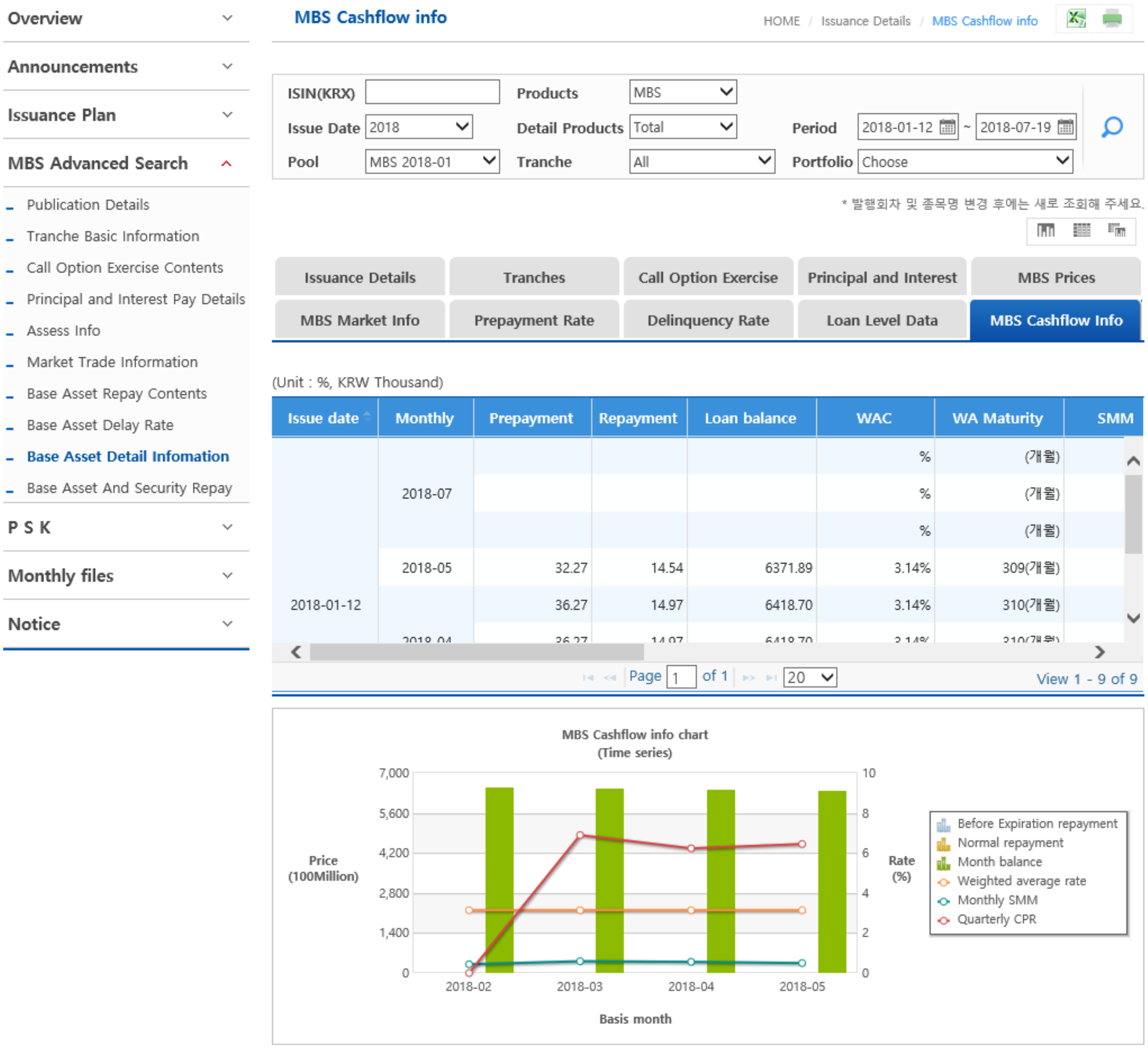


Table C. Data description of Annual pools

		Issue amount KRW1million	Total loan balance, initial KRW1million	Total loan balance, current KRW1million	WARM (month)	Number of loans	Loan name	WAIR, initial (%)	WAIR, current (%)	WAC (%)	Origination spread (%)	CPR	SMM (%)
KMBS04	Avg.	430,867	422,595	745	20.0	6,116	t/e Bogeumjari	6.45	1.90	4.44	2.00	0.26	2.57
(# of pools=7)	SD	95,146	91,898	1,281	34.2	1,146		0.26	3.25	0.41	0.26	0.11	1.43
5pools ended	Min	315,010	309,948	0	0.0	4,546		5.99	0.00	4.06	1.64	0.00	0.00
	Max	570,010	550,507	2,855	70.0	7,631		6.66	6.67	5.00	2.27	0.47	11.22
KMBS05	Avg.	429,010	434,088	3,914	79.9	5,727	t/e Bogeumjari	6.05	6.05	5.04	1.01	0.25	2.54
(# of pools=9)	SD	47,476	42,937	737	2.9	541		0.15	0.17	0.46	0.37	0.12	1.65
	Min	331,010	344,301	2,531	76.0	4,636		5.88	5.85	4.41	0.38	0.00	0.00
	Max	484,010	488,172	4,974	84.0	6,643		6.25	6.26	5.87	1.52	0.53	13.59
KMBS06	Avg.	350,624	351,827	2,925	100.8	4,869	t/e Bogeumjari	6.41	6.43	5.37	1.04	0.27	2.79
(# of pools=5)	SD	32,997	35,594	741	13.3	498		0.16	0.17	0.31	0.27	0.13	1.86
	Min	312,510	310,951	1,807	83.0	4,167		6.25	6.26	5.02	0.61	0.00	0.00
	Max	389,520	390,636	3,676	111.0	5,490		6.61	6.64	5.72	1.27	0.54	13.86
KMBS07	Avg.	364,733	369,234	2,201	127.5	4,513	t/e Bogeumjari	6.15	6.21	5.54	0.61	0.31	3.34
(# of pools=6)	SD	22,622	24,167	476	16.4	377		0.09	0.08	0.28	0.26	0.17	2.65
	Min	337,400	344,806	1,525	103.0	3,848		6.05	6.10	5.17	0.34	0.00	0.00
	Max	392,100	400,866	2,731	150.0	4,888		6.29	6.30	5.86	1.03	0.65	16.84
KMBS08	Avg.	378,286	380,679	1,844	143.0	4,309	t/e Bogeumjari	6.56	6.56	5.92	0.63	0.35	3.85
(# of pools=7)	SD	43,780	43,013	585	14.9	734		0.16	0.15	0.40	0.29	0.18	3.04
	Min	302,501	308,489	1,108	123.0	3,003		6.31	6.39	5.46	0.20	0.00	0.00
	Max	427,300	431,627	2,637	168.0	5,311		6.84	6.85	6.64	1.05	0.67	19.55
KMBS09	Avg.	429,987	423,407	2,780	151.5	4,837	t/e Bogeumjari	6.68	6.79	5.53	1.16	0.39	4.29
(# of pools=8)	SD	102,638	103,227	1,355	9.7	1,343		0.37	0.38	0.17	0.46	0.15	2.78
	Min	313,300	302,944	1,720	137.0	3,267		6.09	6.15	5.22	0.56	0.00	0.00
	Max	554,300	543,736,	5,273	162.0	6,726		7.22	7.34	5.71	1.90	0.71	17.70
KMBS10	Avg.	407,761	399,429	4,802	150.3	4,389	t/e, u Bogeumjari	5.93	5.98	4.78	1.16	0.37	4.23
(# of pools=13)	SD	71,452	69,932	2,760	13.8	583		0.24	0.27	0.37	0.29	0.18	3.06
	Min	308,400	300,241	1,915	129.0	3,618		5.37	5.35	4.29	0.74	0.00	0.00
	Max	563,000	547,319	13,270	169.0	5,752		6.13	6.21	5.39	1.75	0.71	22.82
KMBS11	Avg.	435,626	425,712	14,547	149.7	4,381	t/e, u Bogeumjari	5.38	5.31	4.30	1.08	0.33	3.67
(# of pools=15)	SD	145,717	142,947	6,876	19.7	1,238		0.26	0.42	0.25	0.29	0.17	2.49
	Min	217,300	211,195	4,123	122.0	2,551		5.09	4.30	4.03	0.50	0.00	0.00
	Max	733,800	714,516	33,235	192.0	6,898		5.80	5.81	4.78	1.60	0.64	14.56
KMBS12	Avg.	551,539	539,241	34,284	152.7	5,428	t/e, u Bogeumjari,	4.61	4.58	3.44	1.17	0.32	3.61
(# of pools=33)	SD	211,217	208,792	17,071	38.7	2,489	Conforming	0.31	0.28	0.34	0.16	0.18	2.76
	Min	228,400	221,672	9,575	91.0	2,198		4.19	4.20	3.08	0.91	0.00	0.01
	Max	1,018,800	1,005,714	82,348	224.0	12,729		5.51	5.54	3.99	1.54	0.74	23.12

KMBS13	**Avg.**	681,909	674,933	**89,440**	**155.0**	6,706	**t/e, u Bogeumjari,**	**3.98**	**3.92**	**3.25**	**0.73**	**0.28**	**3.04**
(# of pools=31)	**SD**	210,412	208,323	**37,483**	**39.9**	2,237	**Conforming**	**0.20**	**0.24**	**0.25**	**0.37**	**0.16**	**2.12**
	Min	305,400	303,205	**35,244**	**113.0**	2,739		**3.69**	**3.57**	**2.88**	**0.17**	**0.00**	**0.04**
	Max	1,123,800	1,129,452	**201,887**	**231.0**	12,791		**4.62**	**4.54**	**3.62**	**1.17**	**0.64**	**13.98**
KMBS14	**Avg.**	769,600	758,094	**261,898**	**215.4**	8,152	**u Bogeumjari,**	**3.74**	**3.63**	**3.05**	**0.69**	**0.24**	**2.50**
(# of pools=18)	**SD**	499,323	489,487	**240,241**	**35.9**	5,309	**Didimdol,**	**0.34**	**0.34**	**0.45**	**0.20**	**0.13**	**1.55**
	Min	210,600	208,617	**35,422**	**157.0**	2,407	**Conforming**	**3.31**	**3.26**	**2.37**	**0.32**	**0.00**	**0.11**
	Max	1,766,600	1,738,281	**758,607**	**261.0**	18,885		**4.36**	**4.32**	**3.63**	**0.99**	**0.47**	**10.53**
KMBS15	**Avg.**	2,106,626	2,113,727	**1,336,837**	**249.4**	21,963	**t+, u Bogeumjari,**	**2.96**	**2.95**	**2.27**	**0.70**	**0.10**	**1.08**
(# of pools=26)	**SD**	1,450,850	1,481,119	**1,017,260**	**16.8**	15,347	**Didimdol,**	**0.25**	**0.24**	**0.19**	**0.39**	**0.06**	**0.46**
	Min	403,500	399,185	**268,382**	**212.0**	3,958	**Ansim Conversion,**	**2.63**	**2.64**	**1.97**	**0.12**	**0.00**	**0.12**
	Max	5,062,800	4,974,844	**2,888,629**	**271.0**	51,441	**Conforming**	**3.32**	**3.27**	**2.53**	**1.25**	**0.23**	**3.79**
KMBS16	**Avg.**	1,280,440	1,262,011	**982,144**	**265.0**	11,485	**t+, u Bogeumjari,**	**2.77**	**2.76**	**1.90**	**0.87**	**0.07**	**0.80**
(# of pools=27)	**SD**	415,603	406,709	**326,247**	**21.7**	3,790	**Didimdol,**	**0.18**	**0.17**	**0.27**	**0.37**	**0.04**	**0.27**
	Min	305,100	310,000	**252,110**	**206.0**	3,523	**Conforming**	**2.28**	**2.28**	**1.51**	**-0.01**	**0.00**	**0.08**
	Max	1,789,300	1,750,252	**1,399,664**	**291.0**	17,046		**3.04**	**3.02**	**2.47**	**1.32**	**0.15**	**2.13**
Note.													