ISSN: 2288-7709
JEMM website: https://accesson.kr/jemm
doi: http://dx.doi.org/10.20482/jemm.2025.13.2.73

Predicting MBS Early Prepayment Rates Under External Shocks Using Machine Learning
- Global Financial Crisis vs. COVID-19

Chengai WU[1]

Received: February 24, 2025. Revised: March 15, 2025. Accepted: April 10, 2025.

Abstract

Purpose: This study aims to predict monthly prepayment rates of Mortgage-Backed Securities (MBS) issued by the Korea Housing Finance Corporation, focusing on the effects of external shocks such as the financial crisis and COVID-19. **Research design:** The research compares traditional fixed-effects regression models with machine learning techniques (ElasticNet, LASSO, Ridge) to determine which model best predicts MBS prepayment rates before and after external shocks. **Data and methodology:** The study uses monthly data from June 2004 to December 2020, analyzing MBS prepayment rates alongside various macroeconomic variables. The performance of each model is assessed using cross-validation and blocked cross-validation methods to evaluate stability under different economic conditions. **Results:** Machine learning models, particularly ElasticNet, consistently outperform traditional regression models. ElasticNet showed the highest predictive accuracy, with a stable performance even after the financial crisis and COVID-19, unlike traditional models that struggled to adapt to the shocks. **Conclusions:** The study concludes that machine learning models, especially ElasticNet, offer superior predictive performance in forecasting MBS prepayment rates, especially in volatile market conditions, and should be considered over traditional models for financial predictions.

Keywords: MBS(Mortgage-Backed Securities), Prepayment Rates, Machine Learning, Global Financial Crisis, COVID-19.

JEL Classification Code: E44, F31, F37, G15.

1. Introduction

The financial market has historically experienced significant changes due to various external factors. In particular, the 2008 financial crisis and the 2020 COVID-19 pandemic had profound impacts on the structure and stability of the financial system. These events directly and indirectly affected the real estate market and macroeconomic conditions, influencing securitized products such as Mortgage-Backed Securities (MBS), which are issued based on mortgage loans.

MBS is a type of Asset-Backed Security (ABS) issued based on mortgage loan receivables. In Korea, MBS are currently issued with principal and interest payment guarantees provided by the Korea Housing Finance Corporation, which results in minimal credit risk. However, due to the embedded prepayment option granted to borrowers, there is inherent uncertainty in cash flows.

Consequently, among the primary risks associated with MBS—default risk and prepayment risk—the prepayment risk is considered the more significant factor.

Predicting MBS prepayment rates is vital for investors, helping them to adjust portfolios, optimize strategies, and manage risk. For financial institutions, accurate forecasts

1 First Author. Ph.D. Student, Department of Business Administration, Kyung Hee University, wuchengai89@naver.com.

enhance asset-liability matching and risk management. To accurately assess the value of MBS, it is essential to analyze cash flows and effectively predict prepayment rates.

MBS prepayment rates have traditionally been predicted using statistical models, but these methods have limitations in complex financial markets. With the rise of big data, research applying machine learning and deep learning to financial forecasting has increased globally. While traditional models are still widely used, machine learning has greatly improved prediction accuracy and the analysis of complex data patterns.

For instance, Chakraborty and Joseph (2017) demonstrated that machine learning models achieved high performance in the Bank of England's financial policy and economic forecasting. Similarly, Jagannathan and Bao (2022) found that machine learning models outperformed traditional models in predicting MBS prepayment rates. Furthermore, Chun et al. (2024) revealed that machine learning-based global portfolios using LASSO, Ridge, and ElasticNet models outperformed traditional portfolios in forecasting the KRW/USD exchange rate and stock market trends. These studies highlight the strength of machine learning in effectively modeling complex relationships and improving predictive accuracy.

In South Korea, An et al. (2020) applied LASSO, Ridge, and ElasticNet techniques to predict MBS prepayment rates, identifying ElasticNet as the most suitable model. Yi (2021) also employed machine learning techniques to analyze the impact of Busan's strategic industries on the regional economy. These studies demonstrate the increasing application of machine learning not only in finance but also in policy research.

Under normal economic conditions, the difference in predictive performance between traditional and machine learning models may be negligible. However, following external shocks such as the financial crisis or the COVID-19 pandemic, predictive accuracy can differ significantly.

Such external shocks alter market structures and data patterns, making it essential to understand the interactions between internal variables and external factors. Accurate predictions during economic shocks have practical implications for various market participants, including investors, financial institutions, and governments.

While traditional models, which rely on fixed economic indicators, may fail to adapt effectively to volatility, machine learning models can leverage vast amounts of data to learn complex relationships and patterns, potentially offering superior predictive performance.

This study aims to explore monthly prepayment rate prediction models for MBS issued by the Korea Housing Finance Corporation, using monthly data from June 2004 to December 2020. For this purpose, machine learning techniques, specifically linear regression methods such as LASSO, Ridge, and ElasticNet, will be employed.

The study focuses on two core objectives.

First, it compares the predictive performance of traditional and machine learning models under the impact of external shocks, specifically the financial crisis and the COVID-19 pandemic to determine which model is most suitable. Second, it evaluates predictive performance using both standard K-fold cross-validation and blocked cross-validation methods, which account for time dependencies, to assess which approach is more appropriate for time-series data.

Through this analysis, the study seeks to propose effective approaches for stable predictions following external shocks, contributing to minimizing uncertainty and risk in financial markets. In doing so, it aims to provide practical information to market participants for informed decision-making. Furthermore, the findings can serve as a foundational resource for drawing policy implications aimed at enhancing the stability and efficiency of the MBS market in the future.

The structure of this study is as follows: Chapter 1 presents the introduction, outlining the research background and objectives. Chapter 2 reviews relevant literature. Chapter 3 describes the analytical data used in the study, while Chapter 4 details the research methodology. Chapter 5 presents and compares the results of the empirical analysis. Finally, Chapter 6 offers conclusions, implications, and discusses the study's limitations.

2. Literature Review

Accurately predicting MBS prepayment rates is essential for investors and financial institutions to efficiently manage MBS cash flows, maximize profitability, and optimize portfolio risk. To achieve this, ongoing research focuses on analyzing the factors that influence mortgage prepayment patterns and developing predictive models that reflect the characteristics of the underlying assets.

Existing research on MBS prepayment rates has provided various insights into the factors influencing prepayment behavior. Richard and Roll (1989) analyzed how prepayment rates are determined by the interaction of multiple factors, identifying refinancing incentives, seasoning, seasonality, and burnout effects as significant variables. Similarly, Schwartz and Torous (1989) emphasized that larger interest rate differentials increase the likelihood of prepayment, highlighting the impact of interest rate volatility on prepayment behavior.

Building on Richard and Roll's work, Spahr and Sunderman (1992) argued that refinancing incentives are better explained by the ratio of the contract interest rate to the current interest rate, rather than simply by the absolute interest rate differential. Collectively, these studies underscore the importance of interest rate factors in determining prepayment rates.

In the Korean context, Yoo (2004) employed a Vector Autoregression (VAR) model to analyze the key determinants of MBS prepayment rates. He found that the interest rate spread between market rates (weighted average household loan rates of deposit banks as reported by the Bank of Korea) and contract rates (initial mortgage rates) significantly influenced prepayment behavior. Additionally, he suggested that short-term fluctuations in interest rate spreads could negatively affect prepayments, while housing price volatility had a stronger influence on prepayment rates than broader macroeconomic variables.

Park et al. (2011) conducted a panel regression analysis to identify factors affecting prepayment rates. Their findings indicated that the interest rate spread, loan seasoning, housing price growth rate, and housing transaction volume all had a positive relationship with prepayment rates. These results align with previous studies, reinforcing the significance of interest rate differentials and real estate market fluctuations in determining prepayment behavior.

Jeon et al. (2011) applied multiple regression analysis to examine various factors influencing prepayment rates. Their study highlighted the effects of refinancing incentives (the difference between current mortgage rates and the original contract rate), mortgage seasoning (loan pool age), burnout effects (the ratio of remaining loan balance to initial loan amount), housing price changes, seasonal factors, and the unemployment rate. The results showed positive relationships between prepayment rates and refinancing incentives, loan seasoning, and seasonality, while the unemployment rate had a negative effect.

Choi et al. (2011) analyzed MBS data from 2004 to 2011 using a VAR model, concluding that a decline in market interest rates led to increased refinancing demand, while rising housing prices and higher default rates encouraged prepayments. However, the KOSPI index was found to have no significant relationship with prepayment rates.

Park et al. (2013) utilized a Vector Error Correction Model (VECM) to analyze MBS data from 2004 to 2012. Their results revealed a negative relationship between market interest rates and prepayment rates, while the interest rate spread (the difference between weighted average contract rates and mortgage rates), apartment auction disposal rates, construction industry BSI index, and unemployment rate all showed positive relationships with prepayment rates. Notably, the positive association between the unemployment rate and prepayment rates suggests that prepayment behavior may move counter to broader macroeconomic conditions.

Han et al. (2015) employed a VAR model to examine the influence of mortgage rates on prepayment rates. The study incorporated short- and long-term interest rate differentials, including the CD91 rate, but found that short-term interest rates were not significant predictors. Furthermore, the apartment price index for the Seoul metropolitan area and five major cities exhibited a positive relationship with prepayment rates, while macroeconomic variables such as the KOSPI index and economic growth rate were not found to have significant effects.

Deng et al. (2009) analyzed the Chinese MBS market and found that, like the U.S., the interest rate spread had a positive relationship with prepayment rates. However, macroeconomic variables such as the unemployment rate and stock index significantly influenced China's prepayment rates. Specifically, an increase in the unemployment rate led to higher prepayment rates, whereas the stock index had a negative relationship with prepayment behavior.

Lee et al. (2020) argued that Korea's prepayment rates, like those in China, are influenced by macroeconomic variables. His findings indicated a negative relationship between the consumer price index (CPI) change rate and prepayment rates, while the construction industry BSI index exhibited a positive relationship. This highlights the importance of considering broader macroeconomic conditions, beyond simple interest rate factors, in determining prepayment rates.

With the rapid expansion of big data applications, research incorporating machine learning and deep learning algorithms has increased across various fields. In data-driven research areas such as finance, healthcare, and social sciences, these techniques have demonstrated higher predictive performance than traditional analytical methods, drawing increasing academic and practical interest.

Kleinberg et al. (2015) emphasized that when prediction is prioritized over causal inference, machine learning methods can outperform traditional regression models. Taiyo Saito (2018) applied machine learning techniques (Random Forest, neural networks) to overcome the limitations of traditional logistic regression and survival analysis in predicting MBS prepayment rates, demonstrating superior predictive performance and higher accuracy compared to logistic regression.

In South Korea, research utilizing machine learning techniques has also expanded across various fields. An et al. (2020) analyzed the predictive performance of machine learning algorithms for MBS prepayment rates, finding that ElasticNet demonstrated the highest accuracy. Their study also confirmed that including underlying asset variables, such as LTV and DTI, improved prediction outcomes.

Son (2021) compared machine learning models—Random Forest, GBM, and XGBoost—with traditional logistic regression in predicting mortgage credit risk, revealing that XGBoost had the best performance and suggesting the potential for using machine learning as a supplementary indicator for preferential rates. An et al. (2023) compared traditional linear regression models with machine learning algorithms to assess corporate bond issuance rate prediction performance, demonstrating that machine learning outperformed linear regression and that the optimal model may vary depending on the data characteristics.

3. Research Methods and Materials

3.1. Data and Materials

This study utilized a dataset comprising 23,101 monthly observations across 337 MBS pools issued by the Korea Housing Finance Corporation from June 2004 to December 2020.

To analyze the effects of the financial crisis and the COVID-19 pandemic, the entire dataset was divided into an analytical sample and a predictive sample. The reason for splitting the dataset was to first develop machine learning algorithms using the analytical sample and then evaluate whether these algorithms could accurately predict the monthly prepayment rates using the predictive sample.

The detailed data on monthly MBS prepayment rates and MBS-specific characteristics (such as loan age, loan maturity, contract rate, loan-to-value ratio, debt-to-income ratio, current loan balance, and initial loan amount) were collected from the Korea Housing Finance Corporation's MBS disclosure portal (K-MBS) and subsequently used for analysis. Base rates and market rates were obtained from the Bank of Korea's Economic Statistics System (ECOS). The apartment price index was sourced from the KB Real Estate platform, while macroeconomic variables such as the consumer price index (CPI), KOSPI index, unemployment rate, and economic growth rate (real GDP growth rate) were collected from the National Statistics Portal. The Construction Business Survey Index (BSI) was obtained from the Korea Research Institute for Construction Policy.

Spread represents the value obtained by subtracting the market interest rate (MR) from the contract interest rate (CR), with the contract rate referring to the average contractual interest rate of the underlying MBS assets. The burnout effect was measured as the ratio of the current loan balance to the initial loan amount. Given that the Korea Housing Finance Corporation typically securitizes mortgage loans accumulated over a three-month period, the actual loan age was estimated by adding three months to the pool's elapsed period. A seasonal moving dummy variable was applied, where months characterized by high moving activity (April, May, October, and November) were assigned a value of 1, and all other months were assigned a value of 0.

In the actual analysis phase, all variables were tested for stationarity before being included in the model.

Except for the dependent variable—monthly MBS prepayment rates—all explanatory variables were transformed to align with the analytical objectives. Lagged values (t-1) of explanatory variables were utilized in the model. For macroeconomic variables expected to influence MBS prepayment rates (such as the apartment price index, consumer price index, KOSPI index, and Construction Business Survey Index), the natural logarithm of their lagged values was used as explanatory variables. <Table 1> provides detailed descriptions of the variables used in this study.

Table 1: Summary of Variable Names and Calculation Methods

Category	Variables	Notation	Definition and Calculation Method
Dependent Variable	Monthly Prepayment Rate(%)	SMM	Prepaid amount before maturity / Loan balance at the end of the month
Explanatory Variables	Base Rate(%)	BR	Bank of Korea's base interest rate
	Contract Rate(%)	CR	The weighted average interest rate of underlying assets
	Market Rate(%)	MR	Interest rate on newly issued loans
	Interest Rate Spread(%)	Spread	Contract rate - Market rate
	Loan Age(months)	Age	Pool age + 3 months
	Burnout Effect(%)	BE	Current loan balance / Initial loan amount
	Moving Season Dummy	D_Move	1 if April, May, October, or November; otherwise 0
	Loan Maturity (months)	Matur	Months from loan origination to maturity
	Loan-to-Value Ratio(%)	LTV	Monthly average loan amount / Collateral value
	Debt-to-Income Ratio(%)	DTI	Monthly average total debt repayment amount / Total income
	Apartment Price Index	HPI	Monthly nationwide apartment price index
	Consumer Price Index	CPI	Monthly consumer price index
	KOSPI Index	KOSPI	Monthly Korea Composite Stock Price Index
	Unemployment Rate(%)	UER	Monthly unemployment rate
	Economic Growth Rate(%)	GDP	Quarterly real GDP growth rate

	Construction Business Survey Index	BSI	Monthly construction business survey index

Note: The Single Monthly Mortality (SMM) rate represents the ratio of the amount collected from mortgage borrowers over one month that exceeds the scheduled principal and interest payments, divided by the outstanding MBS loan balance at the beginning of the month. CPI 2020=100, KOSPI 1980.01.04=100

In previous studies, the Conditional Prepayment Rate (CPR), an annualized measure of MBS prepayment rates, has been predominantly used. However, this study employs the Single Month Mortality Rate (SMM) to predict monthly prepayment rates, as it measures prepayment speed on a monthly basis.

The SMM represents the proportion of mortgage loans that have been prepaid and removed from a financial institution's loan portfolio within a specific month, relative to the scheduled balance at the end of that month.

Specifically, it is defined as the ratio of the prepayment amount—calculated by subtracting the actual balance at the end of the month from the scheduled balance—to the scheduled balance at the end of the month. This can be expressed by Equation (1).

$$SMM = \frac{\text{scheduled balance} - \text{actual balance}}{\text{scheduled balance}} \quad (1)$$

3.2. Research Model

3.2.1. Fixed Effects Panel Regression Model

In this study, a fixed-effects panel regression model was selected as the traditional model. To capture the causal relationship between the explanatory variables and the dependent variable, the explanatory variables were lagged.

By one period (t-1). The final panel linear regression model is represented by Equation (2).

$$\begin{aligned} SMM_{i,t} = \alpha &+ \beta_1 BR_{t-1} + \beta_2 CR_{i,t-1} + \beta_3 MR_{t-1} + \beta_4 Spread_{i,t-1} \\ &+ \beta_5 Age_{i,t-1} + \beta_6 BE_{i,t-1} + \beta_7 {D_{Moving}}_{i,t-1} \\ &+ \beta_8 Maturity_{i,t-1} + \beta_9 LTV_{i,t-1} + \beta_{10} DTI_{i,t-1} \\ &+ \beta_{11} HPI_{t-1} + \beta_{12} CPI_{t-1} + \beta_{13} KOSPI_{t-1} \\ &+ \beta_{14} UER_{t-1} + \beta_{15} GDP_{t-1} + \beta_{16} BSI_{t-1} \\ &+ \varepsilon_{i,t} \qquad (2) \end{aligned}$$

3.2.1. LASSO, Ridge, ElasticNet Model

This study employs linear model-based machine learning algorithms Ridge, LASSO, and ElasticNet to predict the continuous variable of monthly MBS prepayment rates. Additionally, the predictive performance of these machine learning models is compared with that of a traditional linear regression model to identify significant explanatory variables.

The study also evaluates the predictive performance of both existing and newly introduced variables, with the ultimate goal of determining the machine learning algorithm that demonstrates the highest predictive accuracy.

The research focuses on utilizing machine learning algorithms that aim to find coefficient estimates (β) that minimize the loss function.

Equations (3), (4), and (5) represent the loss functions of LASSO, Ridge, and ElasticNet, respectively. The regularization terms in each equation denote the regularizer functions incorporated into the loss functions, serving as constraints to minimize in-sample loss.

While adding a regularization function to the loss function may introduce bias in the coefficient estimates within the sample, it helps reduce the variance of the loss across different samples (Hastie, Tibshirani, and Friedman, 2009).

This trade-off allows for a reduction in out-of-sample loss, even at the expense of some in-sample loss. In other words, linear model-based machine learning algorithms address the issue of overfitting by imposing constraints on coefficient values, thereby enhancing predictive performance on new data.

The constraints imposed by the regularization function act to reduce the absolute values of the coefficient estimates, effectively setting the coefficients of variables that do not significantly contribute to prediction to zero. This process also serves as a form of variable selection. The form of the regularization function and thus the distinction between LASSO, Ridge, and ElasticNet models is determined by the parameter values (λ, α) presented in the second term of the equations.

LASSO Regression Loss Function:

$$L_{LASSO}(\beta) = \sum_{i=1}^{n} \left(y_i - x_i^{\top}\beta\right)^2 + \lambda \sum_{j=1}^{p} |\beta_j| \quad (3)$$

In the LASSO regression loss function, $\beta_1,\ \beta_2, \dots,\ \beta_p$ represents the vector of regression coefficients, and λ controls the model complexity by determining the strength of regularization.

The first term, $\sum_{i=1}^{n}\left(y_i - x_i^{\top}\beta\right)^2$ represents the residual sum of squares (RSS), while the second term, $\lambda\sum_{j=1}^{p}|\beta_j|$ corresponds to the L1 regularization term. L1 regularization can shrink some coefficients to 0, effectively performing variable selection.

Ridge Regression Loss Function:

$$L_{Ridge}(\beta) = \sum_{i=1}^{n}\left(y_i - x_i^{\top}\beta\right)^2 + \lambda\sum_{j=1}^{p}\beta_j^{\,2} \quad (4)$$

In this context, β=(β1,β2,…,βp) denotes the vector of regression coefficients, while λ serves as the regularization parameter that controls model complexity. The first term, $\sum_{i=1}^{n}\left(y_i - x_i^{\top}\beta\right)^2$ represents the residual sum of squares (RSS). The second term, $\lambda\sum_{j=1}^{p}\beta_j^{\,2}$ corresponds to the L2 regularization term.

ElasticNet Regression Loss Function:

$$L_{ElasticNet}(\beta) = \sum_{i=1}^{n}\left(y_i - x_i^{\top}\beta\right)^2 + \lambda\left[\alpha\sum_{j=1}^{p}|\beta_j| + \frac{1}{2}(1-\alpha)\sum_{j=1}^{p}\beta_j^{\,2}\right] \quad (5)$$

In this context, β=(β1,β2,…,βp) represents the vector of regression coefficients, while λ and α control the model complexity. The first term, $\sum_{i=1}^{n}\left(y_i - x_i^{\top}\beta\right)^2$ represents the residual sum of squares (RSS). The second term, $\lambda\left[\alpha\sum_{j=1}^{p}|\beta_j| + \frac{1}{2}(1-\alpha)\sum_{j=1}^{p}\beta_j^{\,2}\right]$ is the regularization term, which combines both L1 and L2 regularization.

The key difference between the Ridge and LASSO models lies in the form of constraints imposed on the coefficient values.

Ridge regression employs L2 regularization, which penalizes the squared sum of the coefficients, assigning greater weights to variables that are significant for prediction. This shrinkage reduces the influence of less important variables on new data, thereby enhancing predictive performance.

However, Ridge regression does not set any coefficients exactly to zero, meaning that all explanatory variables remain in the model. This characteristic is particularly beneficial when explanatory variables are highly correlated or when their effects on the dependent variable are similar, as it can improve predictive accuracy.

Nevertheless, a notable drawback arises when dealing with high-dimensional data, as Ridge regression cannot eliminate irrelevant variables, potentially complicating interpretability.

In contrast, the LASSO model was developed to address the limitations of Ridge regression by applying L1 regularization, which penalizes the absolute sum of the coefficients. L1 regularization increases the likelihood that some coefficient estimates will shrink toward zero more aggressively than in Ridge regression, effectively eliminating irrelevant variables by setting their coefficients to zero. This feature allows LASSO to perform variable selection, distinguishing it from Ridge regression.

LASSO demonstrates high predictive performance when the effects of explanatory variables on the dependent variable are concentrated in a subset of variables. By excluding unnecessary variables and focusing only on those critical for prediction, LASSO enhances out-of-sample predictive accuracy.

However, by selecting only a subset of variables, LASSO may risk information loss, potentially reducing predictive performance. This limitation is particularly pronounced when the explanatory variables are highly correlated or exert similar effects on the dependent variable, as LASSO may select only a portion of relevant variables, thereby restricting predictive improvements.

Finally, the ElasticNet model is a hybrid that combines the features of both Ridge and LASSO regression. As shown in Equation (5), the degree to which L1 and L2 norms are applied depends on the value of α. When α=1, ElasticNet becomes equivalent to LASSO regression, while α=0 corresponds to Ridge regression. ElasticNet forms and optimizes a loss function by taking a weighted average of both types of regularization.

This model groups explanatory variables based on their correlations and selects all variables within a group that shares a strong relationship with the dependent variable. Thus, ElasticNet effectively combines the strengths of both Ridge and LASSO. However, in cases where the dataset is small and the number of explanatory variables is limited, ElasticNet may underperform compared to Ridge or LASSO. Additionally, it has the drawback of higher computational costs.

Given that the optimal predictive performance may vary depending on the dataset, it is essential to consider multiple models rather than relying on a single model when constructing prediction algorithms.

The Ridge, LASSO, and ElasticNet algorithms are fundamentally designed to select the optimal λ by minimizing the objective function based on the Mean Squared Error (MSE). The MSE is defined as follows in Equation (6), where n denotes the number of observations, y_i represents the actual value, and $\widehat{y}_i$ denotes the predicted value.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \widehat{y}_i\right)^2 \quad (6)$$

4. Study Design

4.1. Data Splitting

In this study, the entire dataset (23,101 observations) was divided into an analytical sample and a predictive sample to evaluate whether the predictive performance of MBS prepayment rate models remained consistent before and after the financial crisis and the COVID-19 pandemic. The predictive sample was set to cover approximately 12 months before and after each economic shock to allow for a clear comparison of changes across the two periods.

In the training set, repeated experiments were conducted to identify the optimal values for λ and α, allowing the model to learn data patterns and derive the most effective combination of variables for prediction. Based on the selected optimal λ and α, the model with the highest predictive performance was chosen. In the test set, the selected optimal model and hyperparameters from the training phase were applied to perform predictions, evaluating how reliably the model could provide results before and after economic shocks.

For the pre- and post-financial crisis analysis, data from June 2004 to August 2007 was designated as the training set, allowing the model to learn the general patterns of prepayment rates. Subsequently, data from September 2007 to August 2008, just before the onset of the financial crisis, was used as the pre-crisis test set (Test set_Pre-GFC) to assess the model's predictive performance. Data from September 2008 to August 2009, following the outbreak of the financial crisis, was used as the post-crisis test set (Test set_Post-GFC) to evaluate the model's performance after the economic shock.

For the pre- and post-COVID-19 analysis, the same methodological approach was applied. Data from September 2009 to January 2019 was designated as the training set to allow the model to learn the prepayment rate patterns. Subsequently, data from February 2019 to January 2020 was used as the pre-pandemic test set (Test set_Pre-COVID-19) to evaluate predictive performance prior to the pandemic. Finally, data from February 2020 to December 2020 was designated as the post-pandemic test set (Test set_Post-COVID-19) to analyze changes in the model's predictive performance following the COVID-19 shock.

To ensure a valid comparison of predictive performance after both the financial crisis and the COVID-19 pandemic, the data used for the financial crisis analysis was not reused for the COVID-19 analysis. The detailed composition of the dataset is presented in <Table 2>.

Table 2: Sample Composition

DATA SET	Analytical Sample	Prediction Sample	
		Before	After
GFC (1,360)	2004.06 ~2007.08 (581)	2007.09 ~2008.08 (358)	2008.09 ~2009.08 (421)
COVID-19 (21,741)	2009.09 ~2019.01 (15,756)	2019.02 ~2020.01 (2,996)	2020.02 ~2020.12 (2,989)

Note: ( ) indicates the number of observations.

This study utilized both K-fold cross-validation and blocked cross-validation to analyze the training data. In K-fold cross-validation, the data is divided into K folds; each fold serves as a validation set once, while the remaining K-1 folds are used for training. This process repeats until all folds have been used for validation, enabling performance evaluation through alternating training and validation. This method is efficient for limited data, as it utilizes all points for training and is widely used in existing research. Following James et al. (2013), K=10 was set for the analysis.

For machine learning models, the optimal λ (α for ElasticNet) was determined by minimizing the Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE). After selecting variables, each model's performance was compared, and the one with the lowest error was chosen as optimal. The best model's hyperparameters and variable combinations were then applied to the predictive dataset to validate performance and generate final results.

However, applying K-fold cross-validation to time-series data can present several challenges. Since it randomly partitions the data without considering temporal order, future data could inadvertently be used for training while past data serves as validation. This scenario is unrealistic in practical applications and may lead to overestimation of the model's predictive performance by leveraging information from the future. Such distortions can lead to biased model evaluations and poor model selection.

To address this, Bergmeir and Benítez (2012) proposed a cross-validation method tailored for time-series data. Due to concerns regarding time dependence, they recommended using a blocked cross-validation approach for time-series datasets. Similarly, Cerqueira et al. (2020) supported this approach and provided additional evidence that blocked cross-validation can be applied to stationary time-series data.

Given the time-series characteristics of the data used in this study, both standard K-fold cross-validation and blocked cross-validation were considered, with the latter applied to the training data. In blocked cross-validation, each validation set is always preceded by the training set to respect the temporal order. During each iteration, the data is split into non-overlapping blocks, ensuring no overlap between the training and test sets. Each block was designed to include approximately 12 months of data, with 80% allocated for training and 20% for prediction.

In practical applications, where future predictions must be based on past data, blocked cross-validation effectively mimics real-world forecasting scenarios by partitioning the training and test data according to chronological order. <Figure 1> illustrates an example of blocked cross-validation.

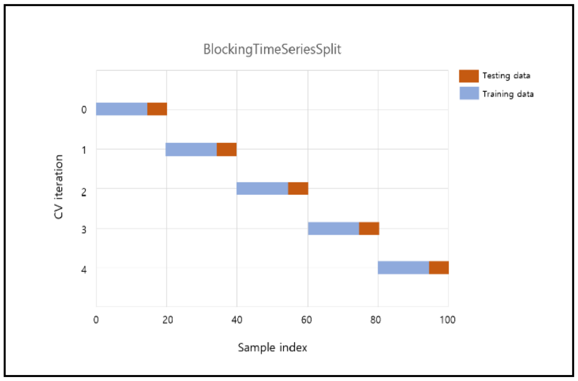

Figure 1: Example of Blocked cross-validation

4.2. Hyperparameter Tuning

Hyperparameter Tuning is the process of adjusting hyperparameter values to optimize the performance of a machine learning model. Hyperparameters are values that must be set before the learning process begins and significantly influence how the model learns from data. Hyperparameter tuning is a critical step in determining how effectively a model can process data and make accurate predictions. For instance, in LASSO and Ridge regression, λ serves as the hyperparameter, while in ElasticNet, both λ and α are treated as hyperparameters.

Several search methods exist for hyperparameter tuning. These include manual search, where values are set based on prior experience; random search, which randomly selects combinations of hyperparameters within a specified range; and Bayesian optimization, which uses probabilistic models to predict function values and selectively explores the most promising regions of the hyperparameter space. However, this study employs the grid search method.

The grid search technique systematically evaluates all possible combinations of hyperparameter values to identify the combination that yields the best performance. This exhaustive search tests every value within a specified range.

The advantage of grid search lies in its thoroughness, as it explores the entire hyperparameter space within the defined range, ensuring that all possible combinations are considered. Additionally, it is intuitive and easy to implement, requiring only the specification of the range and intervals for each hyperparameter.

However, a major drawback of grid search is its computational intensity especially when dealing with a large search space or multiple hyperparameters resulting in increased computational complexity and potentially lengthy processing times.

In this study, the regularization parameter λ is tested over 5,000 values, starting from 0 up to the highest level where all coefficients are reduced to zero. The optimal value of λ is determined based on the highest predictive performance observed. Accordingly, the grid consists of approximately 5,000 values.

Additionally, for the ElasticNet model, α is tuned using grid search within the range of 0 to 1, with intervals of 0.01, allowing for a comparison of predictive performance across different levels of α. <Table 3> presents the tuning ranges of the hyperparameters used for grid search on the training data.

Table 3: Hyperparameter Tuning

Model	Type	Range
LASSO	α	1
	λ	grid: 5,000
Ridge	α	0
	λ	grid: 5,000
ElasticNet	α	$\in$ (0, 1)
		step size: 0.01

4.3. Performance Comparison Metrics

Based on the optimal λ and α values, as well as the selected variables identified from the training set, various models are compared to determine the one with the highest predictive performance. The performance of a machine learning algorithm is determined by the accuracy of its predictions—specifically, the model should exhibit low error rates and strong generalizability.

In machine learning, predictive performance is evaluated by comparing the predicted output values with the actual target values using validation data. This comparison assesses the model's learning capability and overall effectiveness. There are various metrics commonly used to evaluate the predictive power of machine learning models, with the following being the most frequently applied indicators.

4.3.1. RMSE(Root Mean Squared Error)

The Root Mean Squared Error (RMSE) is the square root of the Mean Squared Error (MSE). This transformation addresses the unit issue inherent in MSE, allowing RMSE to share the same units as the original data.

RMSE measures the average magnitude of the difference between predicted and actual values and, like MSE, is particularly sensitive to large errors. It reflects the average deviation between predicted and actual values and is expressed as shown in Equation (7).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (7)$$

4.3.2. MAE(Mean Absolute Error)

The Mean Absolute Error (MAE) is the average of the absolute differences between the predicted and actual values, and it is expressed as shown in Equation (8). Since it does not involve squaring the errors, it is less affected by large deviations. MAE calculates the absolute difference for each data point, assigning equal weight to both large and small errors.

Compared to RMSE, MAE is less sensitive to outliers and large errors, making it a more robust metric in the presence of extreme values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (8)$$

Root Mean Squared Error (RMSE) is more sensitive to large errors, making it advantageous for detecting outliers. In contrast, Mean Absolute Error (MAE) is relatively less sensitive to large errors but is simpler to compute. Since MAE treats all errors equally, regardless of their magnitude, it is particularly useful for assessing the average size of prediction errors.

5. Results

5.1. Descriptive Statistics

An analysis of the descriptive statistics in this study revealed that the dependent variable, the monthly prepayment rate, had a mean of 2.49, with a minimum of 0 and a maximum of 27.28. This indicates that while prepayment rates are generally low, there are instances of sharp increases at certain points. Additionally, explanatory variables were selected with reference to previous research, and the descriptive statistics of the variables are presented in <Table 4>.

Table 4: Descriptive Statistics of Variables

Variables	Obs.	Mean	SD	Min	Max
SMM	23,101	2.49	2.24	0.00	27.28
BR	23,101	1.79	0.90	0.50	5.25
CR	23,101	4.72	1.40	0.09	8.60
MR	23,101	3.46	0.98	2.39	7.58
Spread	23,101	1.26	1.26	-4.23	5.16
Age	23,101	53.58	39.71	1.00	190.00
BE	23,101	2.63	2.51	0.00	37.88
Matur	23,101	191.07	58.59	48.00	324.00
LTV	23,101	59.87	3.13	50.25	67.70
DTI	23,101	32.61	3.82	18.71	51.89
HPI	23,101	64.40	9.23	46.16	82.95
CPI	23,101	95.41	5.31	72.12	100.74
KOSPI	23,101	2,074.60	269.93	735.34	2,873.47
UER	23,101	3.65	0.32	2.90	4.60
GDP	23,101	0.66	0.97	-3.40	3.10
BSI	23,101	75.33	11.97	14.60	101.30

Note: This table presents descriptive statistics. SMM represents the monthly prepayment rate, BR is the base interest rate, CR is the weighted average contract rate of the MBS pool, and MR refers to the market interest rate (newly issued loan rate). Spread represents the value obtained by subtracting the market interest rate from the contract interest rate. Age represents the actual elapsed period and is estimated by adding three months to the pool's elapsed period, as the Korea Housing Finance Corporation typically securitizes pools that are three months old. BE denotes the burnout ratio, calculated as the current loan balance divided by the original loan balance. Matur refers to the loan maturity, measured in months from loan origination to maturity. LTV represents the loan-to-value ratio, and DTI refers to the debt-to-income ratio. HPI indicates the monthly nationwide apartment sales price index, CPI represents the consumer price index, and KOSPI denotes the monthly Korea Composite Stock Price Index. UER refers to the unemployment rate, GDP represents the real growth rate, and BSI indicates the construction industry index.

5.2. Analysis Results

K-fold cross-validation was performed on the training data for both the financial crisis and COVID-19 periods. The optimal parameter values (λ, α) and predictive performance estimates for the traditional fixed-effects model and the machine learning models are presented in <Table 5>.

The analysis using the financial crisis training data revealed that the ElasticNet model demonstrated the best performance among the models tested. The optimal hyperparameters for each model were determined as follows: LASSO (λ=0.00007), Ridge (λ=0.05005), and ElasticNet (λ=0.00005, α=0.98).

The ElasticNet model achieved the lowest error rates, with an RMSE of 0.00326 and an MAE of 0.00255. While LASSO (0.00327, 0.00255) and Ridge (0.00327, 0.00256) exhibited similar performances, ElasticNet produced slightly superior results.

In contrast, the traditional fixed effects(FE) model recorded an RMSE of 0.00339 and an MAE of 0.00267, demonstrating lower predictive performance compared to the machine learning models.

Applying the best-performing ElasticNet model (λ=0.00005, α=0.98) to the pre- and post-financial crisis prediction data revealed consistent predictive strength before the crisis.

In the pre-crisis test set(Test set_Pre-GFC), the ElasticNet model maintained low error rates, with an RMSE of 0.00247 and an MAE of 0.00195, confirming that the machine learning model continued to perform effectively.

This suggests that during this period, the model effectively captured market patterns and that the relationships among variables remained stable, ensuring high explanatory power.

However, in the post-crisis period (Test set_Post-GFC), the model's predictive accuracy significantly declined, with the RMSE rising to 0.00473 and the MAE to 0.00361.

This deterioration in performance is likely due to structural changes in the market environment caused by factors such as changes in lending standards, interest rate fluctuations, and tightening credit conditions after the financial crisis, which likely altered previously learned patterns and relationships among variables.

In the analysis using the COVID-19 training data, ElasticNet again demonstrated the best performance, with the optimal hyperparameters determined as follows: LASSO (λ=0.00015), Ridge(λ=0.05153), ElasticNet (λ=0.00018, α=0.99).

ElasticNet achieved the lowest error values, with an RMSE of 0.02008 and an MAE of 0.01423. Although LASSO (0.02102, 0.01517) and Ridge (0.02102, 0.01517) showed similar performance levels, ElasticNet outperformed them slightly.

In contrast, the traditional FE model recorded an RMSE of 0.02169 and an MAE of 0.01565, demonstrating lower performance compared to the machine learning models.

Applying the selected optimal model, ElasticNet (λ=0.00018, α=0.99), to the pre- and post-COVID-19 prediction data showed that, unlike the financial crisis, model performance did not deteriorate significantly after the pandemic.

In the pre-pandemic prediction data (Test set_Pre COVID19), ElasticNet achieved an RMSE of 0.01291 and an MAE of 0.00744. In the post-pandemic period (Test set_Post COVID19), it maintained similar levels of accuracy, with an RMSE of 0.01297 and an MAE of 0.00749.

In comparison, while the RMSE increased by 91.5% following the financial crisis, it rose by only 0.5% after the COVID-19 outbreak. This suggests that while the financial crisis led to significant structural changes in the economic system that sharply reduced the model's predictive capacity, the COVID-19 pandemic, despite causing economic shocks, did not induce comparable structural shifts in the financial markets. <Table 5> and <Table 6> present the results of the financial crisis and COVID-19 pre- and post-shock training and prediction analyses conducted using K-fold cross-validation.

Table 5: GFC Training and Prediction Results

Model	Optimal		MSE	RMSE	MAE
	λ	α			
Training set GFC					
LASSO	0.00007		0.00001	0.00327	0.00255
Ridge	0.05005		0.00001	0.00327	0.00256
ElasticNet	0.00005	0.98	0.00001	0.00326	0.00255
FE	-	-	0.00001	0.00339	0.00267
Test set_Pre-GFC					
ElasticNet	0.00005	0.98	0.00001	0.00247	0.00195
Test set _Post-GFC					
ElasticNet	0.00005	0.98	0.00002	0.00473	0.00361

Table 6: COVID19 Training and Prediction Results

Model	Optimal		MSE	RMSE	MAE
	λ	α			
Training set COVID19					
LASSO	0.00015		0.00044	0.02102	0.01517
Ridge	0.05153		0.00044	0.02102	0.01517
ElasticNet	0.00018	0.99	0.00040	0.02008	0.01423
FE	-	-	0.00047	0.02169	0.01565
Test set_Pre-COVID19					
ElasticNet	0.00018	0.99	0.00017	0.01291	0.00744
Test set _Post-COVID19					
ElasticNet	0.00018	0.99	0.00018	0.01297	0.00749

Using blocked cross-validation, the analysis of financial crisis and COVID-19 data showed that machine learning models outperformed the FE model, with ElasticNet achieving the highest accuracy.

In the financial crisis training data, ElasticNet recorded the lowest error values, with an RMSE of 0.00230 and an MAE of 0.00198. While LASSO (0.00315, 0.00268) and Ridge (0.00257, 0.00226) also delivered comparable results, ElasticNet showed slightly superior performance.

Similarly, in the COVID-19 training data, ElasticNet achieved the best results with an RMSE of 0.00235 and an MAE of 0.00190, outperforming both LASSO (0.00377, 0.00342) and Ridge (0.00437, 0.00408). In contrast, the FE model demonstrated significantly lower predictive performance than the machine learning models.

When applying the optimal models derived from the training data to the pre- and post-financial crisis and COVID-19 prediction datasets, it was observed that predictive performance deteriorated more sharply after the financial crisis than after COVID-19.

For the pre-crisis period (Test set_Pre-GFC), ElasticNet maintained strong predictive accuracy with an RMSE of 0.00325 and an MAE of 0.00252. However, after the financial crisis (Test set_Post-GFC), predictive performance declined, with the RMSE increasing to 0.00297 and the

MAE to 0.00228.

Notably, the FE model exhibited a substantial deterioration in performance after the financial crisis, with the RMSE rising dramatically to 0.04096 and the MAE to 0.03510.

In contrast, during the pre-COVID-19 period (Test set_Pre-COVID19), ElasticNet achieved an RMSE of 0.00935 and an MAE of 0.00598, indicating slightly higher predictive accuracy than in the pre-financial crisis period. After the COVID-19 outbreak (Test set_Post-COVID19), the model's performance only slightly deteriorated, with the RMSE increasing to 0.01181 and the MAE to 0.00707.

When evaluating the decline in predictive performance of the FE model, a stark contrast emerges: after the financial crisis, the RMSE surged to 0.04096, whereas after COVID-19, the increase was more moderate, reaching 0.01712.

These results suggest that the financial crisis caused significant structural changes in the market, leading to a substantial decline in the predictive power of existing models.

In contrast, while COVID-19 had economic impacts, it did not induce structural changes in the financial markets to the same extent as the financial crisis, resulting smaller decline in predictive performance. <Table 7> and <Table 8> present the pre- and post-shock training and prediction results for the financial crisis and COVID-19 periods using blocked cross-validation.

Table 7: GFC Training and Prediction Results

Model	Optimal		MSE	RMSE	MAE
	λ	α			
Training set GFC					
LASSO	0.00020		0.00001	0.00315	0.00268
Ridge	0.00371		0.00001	0.00257	0.00226
ElasticNet	0.00001	0.93	0.00001	0.00230	0.00198
FE	-	-	0.00001	0.00377	0.00319
Test set_Pre-GFC					
ElasticNet	0.00001	0.93	0.00001	0.00325	0.00252
FE	-	-	0.00020	0.01420	0.01252
Test set _Post-GFC					
ElasticNet	0.00001	0.93	0.00001	0.00297	0.00228
FE	-	-	0.00168	0.04096	0.03510

Table 8: COVID19 Training and Prediction Results

Model	Optimal		MSE	RMSE	MAE
	λ	α			
Training set COVID19					
LASSO	0.00014		0.00001	0.00377	0.00342
Ridge	0.30660		0.00002	0.00437	0.00408
ElasticNet	0.00005	0.92	0.00001	0.00235	0.00190
FE	-	-	0.00770	0.08777	0.08617
Test set_Pre-COVID19					
ElasticNet	0.00005	0.92	0.00009	0.00935	0.00598
FE	-	-	0.00022	0.01494	0.01121
Test set _Post-COVID19					
ElasticNet	0.00005	0.92	0.00014	0.01181	0.00707
FE	-	-	0.00029	0.01712	0.01216

5.3. Results Comparison

5.3.1. Traditional Regression Models vs. Machine Learning Algorithms

In both the financial crisis and COVID-19 training datasets, machine learning models consistently outperformed the traditional regression model in terms of predictive performance, with ElasticNet demonstrating the highest level of accuracy. This trend remained consistent when applied to the prediction data, where ElasticNet also maintained superior and more stable performance compared to the traditional FE model.

When dealing with complex datasets characterized by high volatility and multicollinearity, as is typical in financial markets, ElasticNet's ability to perform variable selection and regularization appears to contribute significantly to enhancing model performance.

Therefore, ElasticNet offers more reliable predictive accuracy than traditional regression models and serves as an effective modeling choice in environments with high uncertainty, such as financial markets.

5.3.2. Performance under Exogenous Shocks (GFC vs. COVID-19)

The decline in predictive performance was significantly more pronounced after the financial crisis than following the COVID-19 pandemic. After the financial crisis, the FE model experienced a sharp increase in both RMSE and MAE, resulting in a substantial loss of predictive accuracy. Although ElasticNet also exhibited some performance degradation, it was less severe compared to the FE model.

In contrast, following the COVID-19 pandemic, the deterioration in model performance was minimal, and ElasticNet maintained stable predictive accuracy.

This discrepancy appears to stem from the structural changes in the financial markets caused by the financial crisis. Factors such as credit tightening, interest rate fluctuations, and changes in lending standards fundamentally altered market dynamics after the crisis.

On the other hand, while COVID-19 caused temporary volatility, the explanatory power of the existing variables remained largely intact.

After the financial crisis, the previously learned patterns in traditional models became obsolete, leading to a sharp decline in predictive power. However, following COVID-

19, the existing variables continued to hold explanatory relevance, allowing the models to maintain stable predictive performance.

5.3.3. Cross-Validation Methods(K-fold Cross-Validation vs. Blocked Cross-Validation)

K-fold cross-validation has limitations when applied to time-series data, as it randomly partitions the training data without considering temporal dependencies, thus failing to capture causal relationships across time periods.

This limitation is particularly problematic in time-series data, such as that from financial markets, where changes in market conditions can significantly affect model performance. Although predictions may be made using new data, discrepancies can arise between the expected performance observed during training and the actual performance in real-world forecasting, due to the disconnect between the K-fold cross-validation framework and the actual prediction environment.

Consequently, even if a model demonstrates high performance in K-fold cross-validation, its predictive accuracy may deteriorate when applied to new data.

In contrast, blocked cross-validation follows a structure that predicts future outcomes based on past data, thereby reflecting the actual forecasting environment more accurately. This method allows for a more realistic evaluation of model performance, as it mirrors the chronological sequence of real-world predictions.

Therefore, for time-series data, particularly in financial markets, blocked cross-validation is a more reliable model evaluation method than standard K-fold cross-validation. It enables a more accurate assessment of how effectively a model will perform in a real-world predictive setting.

6. Conclusion

This study compares and analyzes fixed-effects regression analysis (traditional model) with machine learning algorithms, including LASSO, Ridge, and ElasticNet, to predict MBS prepayment rates following external shocks such as the financial crisis and the COVID-19 pandemic.

For the machine learning models, in addition to applying standard K-fold cross-validation, blocked cross-validation was employed to account for the time-series characteristics of the data, enabling a more accurate evaluation of model performance.

By utilizing a diverse set of variables, this study aims to identify the optimal model and hyperparameters for each algorithm and to propose an improved prepayment rate prediction algorithm with enhanced predictive performance.

First, machine learning models exhibited, on average, higher predictive accuracy than traditional regression models, with their superiority becoming particularly evident during periods of external shocks, such as the financial crisis and the COVID-19 pandemic.

This advantage stems from the ability of machine learning models to learn patterns and generalize from data.

By automatically identifying complex patterns within large datasets, these models can provide stable predictions even in challenging and unpredictable conditions.

Moreover, by employing various regularization techniques, they effectively prevent overfitting and reduce statistical uncertainty, thereby enhancing predictive accuracy.

A comparative analysis of various machine learning models (LASSO, Ridge, ElasticNet) and the traditional regression model (FE) demonstrated that the ElasticNet model achieved the highest predictive accuracy. Notably, ElasticNet maintained stable predictive performance even under external shock conditions, such as the financial crisis and the COVID-19 pandemic.

By combining L1 and L2 regularization, ElasticNet simultaneously controls model complexity and performs variable selection. This allows the model to effectively manage high-dimensional data and address multicollinearity between variables, capturing essential features of the data and enabling robust, stable predictions.

In conclusion, the empirical validation of the ElasticNet model's superior performance and the generalization capabilities of machine learning techniques highlight the effectiveness of data-driven machine learning methods in predicting MBS prepayment rates.

Particularly when considering robustness in forecasting after external shocks, the use of machine learning-based predictive models offers greater value compared to traditional models, providing a more reliable approach for managing uncertainty in volatile financial markets.

Second, predictive performance after COVID-19 was more favorable compared to the period following the financial crisis. While the model’s predictive accuracy deteriorated significantly after the financial crisis, it remained stable in the post-COVID-19 period.

This can be attributed to the fact that the financial crisis-induced fundamental changes in market structure and capital flows, whereas COVID-19 caused only a temporary shock, with existing variables retaining their explanatory power.

Following the financial crisis, factors such as changes in lending standards, credit tightening, and interest rate fluctuations collectively contributed to the model's inability to adapt to the new market environment.

In contrast, during the COVID-19 pandemic, proactive government stimulus measures and liquidity injections helped the market recover swiftly, allowing predictive performance to remain stable.

This suggests that, unlike the financial crisis, COVID-19 did not induce structural changes in the financial system.

These findings highlight the limitations of traditional regression models in the face of structural market changes, such as those experienced during the financial crisis, and suggest that the application of machine learning models is more effective in such scenarios.

While traditional models can still provide a reasonable level of predictive accuracy in response to short-term shocks, such as COVID-19, more precise forecasting requires machine learning models like ElasticNet.

By leveraging variable selection and regularization, ElasticNet contributes to maintaining stable performance even in the face of temporary market disruptions.

Third, blocked cross-validation, which accounts for time dependencies, proved to be a more realistic method for model evaluation.

In standard K-fold cross-validation, data is randomly partitioned during the training process, which can lead to the unintended inclusion of future data in the training set.

This introduces the risk of overestimating predictive performance due to data leakage. In contrast, blocked cross-validation is structured to predict future data based on past observations, making it a more appropriate method for ensuring realistic predictive accuracy.

This approach is particularly effective in scenarios involving structural changes in the real estate market caused by economic shocks, such as those following the financial crisis.

In such cases, traditional K-fold cross-validation fails to capture the time-series nature of the data, making it difficult to accurately assess the model's generalization capability.

Blocked cross-validation, however, excludes future data from the training process and conducts validation sequentially, enabling evaluations that closely reflect real-world forecasting environments.

This method is particularly valuable for objectively assessing whether existing models can maintain consistent performance in the face of external shocks, such as the financial crisis or the COVID-19 pandemic. By ensuring that predictions are made in a forward-looking manner, blocked cross-validation offers a more accurate reflection of how a model would perform in practical, time-dependent prediction scenarios.

This study empirically confirms the superior performance of machine learning models in predicting MBS prepayment rates following external shocks, such as the financial crisis and the COVID-19 pandemic. Additionally, it demonstrates that blocked cross-validation, which accounts for temporal dependencies, serves as a more reliable evaluation method for time-series data analysis.

These findings suggest that in complex environments, such as financial markets, machine learning-based models can offer more effective predictive performance than traditional regression models. In particular, during periods of market structural change, such as those experienced during financial crises it is crucial to apply machine learning models to achieve more robust and accurate forecasts.

However, this study has several limitations.

First, the variables were primarily constructed around macroeconomic indicators and market interest rates, without adequately incorporating micro-level variables such as borrowers' credit scores, income levels, and loan types. Including such variables could potentially enhance the accuracy of prepayment rate predictions.

Second, while machine learning models were utilized to forecast prepayment rates following the financial crisis and the COVID-19 pandemic, more sophisticated models that combine linear and nonlinear relationships were not considered. Incorporating such advanced models could further improve predictive performance.

Finally, the analysis was conducted using data focused on specific periods, raising the question of whether similar predictive performance would hold under different types of external shocks. Future research should aim to increase the generalizability of the model by utilizing datasets that reflect a wider range of economic shocks.

References

An, J. Y., and Lim, B. K. (2020). The Predictive Power of Machine Learning Algorithm: The Case of MBS Prepayment Rates, *Journal of Money & Finance*, *34*(2), 33-63.

An, J. Y., and Lim, B. K.(2023). Forecasting Corporate Bond Yields with Machine Learning, *Korea Financial Management Association*, *40*(5), 213–243.

Bergmeier, C., Benitez, J.(2012). On the use of cross-validation for time series predictor evaluation. *Information Sciences*, *191*, 192–213.

Cerqueira, V., Torgo, L., Mozetic, I.(2020). Evaluating time series forecasting models: an empirical study on performance estimation methods, *Machine Learning*, *109*(11), 1997–2028.

Chakraborty, C., and Joseph, A., 2017, Machine Learning at Central Banks, *Bank of England Working Paper*, No. 674.

Choi, S. D., and Kim, S. T.(2011). Determinants of Prepayment Risk in MBS: Focusing on MBS Issued by the Korea Housing Finance Corporation, *Journal of Financial Engineering*, *10*(4), 81-103.

Chun, D., Kang, J., and Kim, J.(2024). Forecasting Returns with Machine Learning and Optimizing Global Portfolios: Evidence from the Korean and U.S. Stock Markets, *Financial Innovation*, *10*, 1-30.

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

Han, S. H., Wang, B., Lee, C. S., and Kang, M. G.(2015). The Relationship Between Real Estate Business Cycles and MBS

Prepayment and Its Policy Implications, *Regional Studies*, *31*(4), 91-105.

Hastie, T., R. Tibshirani, and J. Friedman.(2009). The elements of statistical learning: data mining, inference, and prediction, *Springer Science & Business Media*.

Jagannathan, N. A., and Bao, Q.(2022). Machine Learning–Based Systematic Investing in Agency Mortgage-Backed Securities, *The Journal of Financial Data Science*, *4*(4), 1-15.

James, G., D. Witten, T. Hastie, and R. Tibshirani.(2013). An Introduction to Statistical Learning, *New York: Springer*, *112*.

Jeon, J. K., and Kim, C. K.(2011).An Empirical Study on the Mortgage Loan Prepayment Model of Korea, *Korean Journal of Financial Management*, *28*(4), 173-199.

Kim, C. K., and Lim, B. K.(2020).The Effect of Mortgage Prepayment Charges on the MBS Prepayment Rates, *Korean Academic Society of Business Administration*, *49*(3), 717-738.

Kim, H., Cho, H., and Ryu, D.(2018).Characteristics of Mortgage Terminations: An Analysis of a Loan-level Dataset, *The Journal of Real Estate Finance and Economics*, *57*(4), 647-676.

Kim, H. J., Ryu, D. J., and Cho, H.(2018). An Empirical Study on the Prepayment of Mortgages in Korea: Focusing on the Effects of Safe Conversion Loans, *Korean Management Review*, *47*(4), 865-887.

Kleinberg, J., Ludwig, J., Mullainathan, S., and Obermeyer, Z.(2015).Prediction Policy Problems, *American Economic Review*, *105*(5), 491-495.

Lee, H.J., and Lee, J.M.(2020).Factors Affecting Prepayment Risk of Underlying Assets in Domestic MBS, *Residential Environment*, *18*(1), 305-317.

Park, D. K., and Kim, D. H.(2013). A Study on the Factors Affecting Prepayment Risk of Domestic MBS: Focusing on the VEC Model, *Real Estate Studies*, *23*(3), 157-179.

Park, Y. W., and Bang, D. W.(2011). Analysis of Determinants of the Conditional Prepayment Rate of Fixed-Rate Mortgage Loans: A Time-Series Regression Analysis Based on Loan Pools, *Housing Studies Review*, *19*(3), 77-99.

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

Saito, T.(2018) Mortgage Prepayment Rate Estimation with Machine Learning, Master's Thesis, *Delft University of Technology.*

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

Son, J. M.(2021). Optimum Machine Learning Method Predicting Credit Risk in Housing Mortgage Loan, *Korean Institute of Information Scientists and Engineers*, 700–702.

Spahr, R. W., and Sunderman, M. A.(1992).The Effect of Prepayment Modeling in Pricing Mortgage-Backed Securities, *Journal of Housing Research*, 3, 381-400.

Yi, C. D.(2021). Machine Learning Prediction of Economic Effects of Busan's Strategic Industry through Ridge Regression and Lasso Regression, *Journal of Korea Port Economic Association*, *37*(1), 197-215.

Yoon, J. H., and Kim, J. I.(2023).An Empirical Research on Heterogeneity of Mortgage Prepayment Behavior: Focusing on Debtor's Income and Age Level, *Korean Journal of Financial Studies*, *52*(4), 519-542.

Yoo, S. D.(2004). A Study on Prepayment of Mortgage Loans: Focusing on MBS 2000-1, *Korea Real Estate Research Institute*, *14*(2), 139-163.