

# THE ECONOMIC EXPLAINABILITY OF MACHINE LEARNING AND STANDARD ECONOMETRIC MODELS-AN APPLICATION TO THE U.S. MORTGAGE DEFAULT RISK

Dong-sup KIM, Seungwoo SHIN \*

*School of Real Estate, Konkuk University, Seoul, Korea*

Received 10 February 2021; accepted 10 June 2021

**Abstract.** This study aims to bridge the gap between two perspectives of explainability—machine learning and engineering, and economics and standard econometrics—by applying three marginal measurements. The existing real estate literature has primarily used econometric models to analyze the factors that affect the default risk of mortgage loans. However, in this study, we estimate a default risk model using a machine learning-based approach with the help of a U.S. securitized mortgage loan database. Moreover, we compare the economic explainability of the models by calculating the marginal effect and marginal importance of individual risk factors using both econometric and machine learning approaches. Machine learning-based models are quite effective in terms of predictive power; however, the general perception is that they do not efficiently explain the causal relationships within them. This study utilizes the concepts of marginal effects and marginal importance to compare the explanatory power of individual input variables in various models. This can simultaneously help improve the explainability of machine learning techniques and enhance the performance of standard econometric methods.

**Keywords:** mortgage loan, default risk, machine learning, explainability, marginal effect, partial dependence plot, SHAP.

## Introduction

This study uses machine learning (ML) and standard econometric methodologies to analyze the U.S. mortgage default risk data obtained from Freddie Mac. Compared to standard econometric techniques, ML techniques are considered to have greater learning and forecasting capabilities; therefore, in practice, they play an effective role in the detailed management of default risk. However, unlike standard econometric approaches, ML techniques make a limited contribution to explaining the cause of phenomena or being applied in practice while holding accountability. Such techniques are viewed less favorably in the highly regulated field of credit risk management because the complexity and non-linearity of ML models and how their predictions are derived make them difficult for ordinary people to understand (Štrumbelj & Kononenko, 2011). Recently, numerous attempts have been made in the field of artificial intelligence to solve the problem of ML explainability and interpretability. Referred to as explainable artificial intelligence (XAI), this field attempts to produce results that can be understood by humans.

In this context, explainability can be addressed from two different perspectives: ML and engineering, and economics and standard econometrics. While Campbell and Cocco (2015) explain residential mortgage default from the perspective of economics and standard econometrics, Bracke et al. (2019) observe it from an ML and engineering perspective. Although words such as explain, explanation, explainable, and explainability are closely related in terms of their meaning, Campbell and Cocco's (2015) use of such terms has virtually no relationship with their use in Bracke et al. (2019).

This study aims to bridge the gap between the two perspectives of Bracke et al. (2019) and Campbell and Cocco (2015) using economic and econometric methods. First, we describe the current interpretation of explainability in ML models by evaluating the various stakeholders in this field (Preece et al., 2018). Second, we list similar concepts, including explainable, interpretable, understandable, comprehensible, accountable, responsible, justifiable, ethical, transparent, causality, trustworthy, and glass box (Arrieta et al., 2020). Third, we introduce the concept of marginal effect (ME) from an economics perspective, which is based on the rigorous notion of economic causality.

\*Corresponding author. E-mail: [ss244@konkuk.ac.kr](mailto:ss244@konkuk.ac.kr)

Interpretability and explainability are important to stakeholders such as data scientists, end users, and regulators, although for different reasons (Kabul, 2017). Data scientists aim at building models whose predictions are highly accurate (developers and theorists in Preece et al., 2018) by determining the best algorithm to improve a specific model. In this context, the major stakeholders may be the end users who primarily focus on understanding the factors behind a particular model's generation of a certain prediction (expert users, directors of a company, data subjects, and affected customers in Arrieta et al., 2020). These stakeholders want to know how such decisions affect them. Further, end users try to investigate whether they are subject to unbiased treatment and the necessity of objecting to a particular decision. Regulators and lawmakers try to protect end users; hence, regulations have been recently issued, such as the General Data Protection Regulation and EU Expert Group on AI (Bücker et al., 2020), Equal Credit Opportunity Act and Fair Credit Reporting Act (Chen, 2018), and Payment Service Directives 2 (Torrent et al., 2020).

However, with the inevitable rise in complex ensemble ML algorithms, regulators are becoming increasingly concerned about the decisions made by ML models, especially in the field of financial services. Initially, an ensemble model simply referred to a conglomeration of models (Nanni & Lumini, 2009). However, at present, the ensemble approach combines virtually anything from models to feature selection methods to improve prediction performance (Dahiya et al., 2017; Koutanaei et al., 2015). This trend of increasing ensemble modeling has not depended on the types of classifiers. In this context, prior studies have investigated decision tree (Wang et al., 2012) and random forest (RF) ensembles (Chopra & Bhilare, 2018), support vector machine (Pławiak et al., 2019), and neural networks, especially in the fields of credit scoring and credit risk management (Tsai & Wu, 2008). Overall, stakeholders want black box models to exhibit qualities similar to those of other causal models, such as transparency, trustworthiness, and explainability (Shead, 2020).

Analyses in the field of economics and econometrics involve a strong notion of cause-effect relationships, applying a rigorous background of utility (profit or cost) theory (Campbell & Cocco, 2015). In contrast, causality is not among the most important goals for studies related to engineering (Arrieta et al., 2020).

In this study, causality and explainability are described from the perspective of economics and standard econometrics; we explain our model via ME and test for statistical significance. Studies in the field of ML prefer using the term marginal contribution, that is, significant influence or causal importance, to measure explainability.

To understand both the model and its prediction, we estimate the model and calculate the ME of each independent variable on default risk, which is taken as the dependent variable. The explanatory power of ML models can be evaluated using ME analysis. A comparison of the

alternative models can be conducted by implementing the concepts of ME and marginal importance (partial dependence plot [PDP] and Shapley additive explanation value [SHAP]). However, we do not compare their prediction power but focus on three different measures of explainability based on both economics and ML perspectives.

This study uses residential mortgage data comprising 15,904 loans originated from January 2005 to November 2010. We develop a residential mortgage risk prediction model by estimating the probability of default for each loan. To achieve this, we estimate a default risk model by employing logit regression, which is a standard econometric method, as well as the artificial neural network (ANN) and RF methods used in ML.

We choose RF and ANN among other ML approaches because both models are relatively simple compared to other deep learning approaches. If competing models are too different from each other, we might fail to compare them on an apples-to-apples basis. This restriction of similarity also influences the logit model. RF cannot be applied to non-i.i.d. data, inclusive of panel datasets (Pearl, 2019; Steinwart et al., 2009). To feasibly compare these three approaches, we fit a simple logit model instead of a panel logit model and use the information in the event month, where “event” refers to either default or right censored in the last observation month of September 2013, following Bracke et al. (2019).

The contributions of this study are as follows. First, our approach is distinctive in the sense that we try to bridge both approaches by applying ME from the economic perspective. Specifically, we explain the ML technique using the concept of ME, which is more conducive or intuitive for human understanding while also being based on economic theory. ME has many advantages; it is a model agnostic approach, as it does not depend on the model specification. It produces results very quickly, and implementation time does not depend on either the number of variables or the model specification. Moreover, ME does not require either model re-training or re-estimation or both. ME analysis reports the directional sign of the independent variable's influence on the dependent variable. In a measure of explainability, statistical significance is of primary importance (Chernozhukov et al., 2018; Mullainathan & Spiess, 2017); ME provides the useful tool of a statistical significance test. For government regulators, ME offers feasible guidelines regarding the required size (strength) of a policy variable change based on the targeted statistical significance level, such as 1%, 5%, or 10%.

Second, we also utilize the marginal importance measure of the ML model to conduct a detailed supplementary analysis of the standard econometrics model, following Athey and Imbens (2017), Mullainathan and Spiess (2017), Chernozhukov et al. (2018), Rudin (2019), and Torrent et al. (2020). Third, we examine specific non-linear functional forms of input variables, such as squared and cubed polynomials, and test the similarity of the models using the Mann-Whitney test.

Overall, this study utilizes the advantages of two alternative methodologies to propose a better model. By successfully formulating a model, we expect to achieve a more transparent and trustworthy ML model that is based on economic explainability and produces more accurate predictions.

The paper is structured as follows. The introduction explains the background and purpose of the study. In Section 1, we examine prior studies related to the econometrics model and ML approach by considering residential mortgage default and credit scoring models. In Section 2, we describe the measures of explainability. In Section 3, the results of the empirical analyses of the default model through econometrics and ML models are compared using the concepts of ME and marginal importance. We conclude the paper by summarizing the results and acknowledging its limitations.

## 1. Prior research

While the history of the development of basic mortgage modeling is studied by Wallace (2005), Campbell and Cocco (2015) examine its theoretical modeling from an economic perspective. In this context, we review the following related studies in the existing literature.

### 1.1. ML-based mortgage default risk

A majority of the studies that examine ML-based default risk analysis focus on prediction performance rather than logical explanatory power or economic intuition (Feldman & Gross, 2005; Fitzpatrick & Mues, 2016; Fuster et al., 2020; Galindo & Tamayo, 2000; Kvamme et al., 2018; Sealand, 2018; Sirignano et al., 2016).

From the perspective of social science, Feldman and Gross (2005) were the first to use classification and regression trees (CART) in mortgage credit analysis. Sirignano et al. (2016) use a deep learning approach, while Galindo and Tamayo (2000) compare the prediction error of both standard econometric and ML models and find that CART performs the best. Fuster et al. (2020) observe the effect of the change in practice from linear to non-linear estimation models from the perspective of the people affected, especially in terms of race and gender. This change to non-linear models may produce winners and losers and reduce the cross-subsidy resulting from the linear model, which depends on the conditional expectations among multiple groups.

### 1.2. ML-based explainability measures

Explainability measures have played an exceptionally important role in current studies. Here, we analyze only the studies related to credit risk management. Explainability measures include PDP and individual contribution expectation (ICE) (Bücker et al., 2020; Fahner, 2018; Goldstein et al., 2015; Zhao & Hastie, 2021), SHAP (Bücker et al., 2020; Bussmann et al., 2021; Chen, 2018; Datta et al., 2016; Torrent et al., 2020; Lundberg & Lee, 2017), SHAP

flow (Wang et al., 2021), and LIME (Bücker et al., 2020). This study reports only PDP and SHAP measures for simplicity.

To improve the explainability of ML models, Rudin (2019) insists on building a model that can be explained fundamentally, rather than trying to explain a complex model. To this end, the model should reflect a viewpoint that emphasizes both domain knowledge (Chen, 2018; Fahner, 2018; Torrent et al., 2020; Zhao & Hastie, 2021) and the behavioral characteristics of stakeholders (Athey & Imbens, 2017; Mullainathan & Spiess, 2017; Rudin, 2019).

Moreover, existing studies focus on improving the explainability of ML models by imposing structural constraints in terms of monotonicity and linear relationships (Fahner, 2018), sparsity and low dimensionality (Mullainathan & Spiess, 2017; Rudin, 2019), and additive and multiplicative structures (Bussmann et al., 2021; Zhao & Hastie, 2021).

## 2. Methodology

Due to their non-parametric approach and non-linear structure, it is difficult to explain the prediction results of ML techniques; therefore, it is difficult to understand the model's internal structure (Arrieta et al., 2020). Moreover, there are few to no generally accepted definitions of explainability (Preece et al., 2018), similar to the definition of mortgage default itself.

During the process of analyzing explainability, the econometric and ML approaches are separately applied to the dataset to derive the model estimation results. Tools for estimating explainability are applied to the derived model and comparison analysis is discussed.

Specifically, we employ logit regression analysis along with ANN and RF techniques to estimate a default risk model for U.S. residential mortgage loans. To explain the parameters that influence the ML models, the ANN possessed one or two hidden layers, and the number of hidden nodes were either matched with the number of input variables or increased up to three times the number of variables. Fundamentally, one hidden layer was used along with the same number of hidden nodes and input variables to determine whether the explainability of the features applied to the ML-based model can be estimated utilizing econometrics-based ME techniques.

RF is an ensemble technique that makes a final decision through the majority vote of the results that were determined by multiple decision trees. For the prediction result (probability value) to be continuously distributed, a sufficient number of trees must be generated; a minimum of 100 trees was deemed appropriate for this study to allow continuous calculation of the probability value. Therefore, 100 decision trees were utilized.

The following provides a brief introduction to the methods of estimating explainability in ML. First, utilizing the visualization tool, PDP, we graph the changes in the dependent variable that result from a change in the

value of the independent variable(s). Additionally, the SHAP value is estimated to comprehend the influence of individual variable(s).

In addition to methods such as PDP and the SHAP value, this study introduces the concept of ME to estimate the explainability of logit regression analysis and ML models. ME indicates how infinitesimal changes in the independent variable affect the dependent variable. For example, ME has the same meaning as the regression coefficients of each variable in a linear regression model. Due to the absence of regression coefficients in ML, the concept of ME is introduced and utilized operationally. Non-linear model, such as nested logit model, also utilizes ME to calculate the scale of the causal effect of independent variables on dependent variables.

A simple ME approach is applied to perturb single inputs and measure the change in dependent variables in the model, that is, changing one input to observe the outputs. This method estimates the response of the dependent variable in the model to a specific independent variable on a single point as a linear approximation (Lundberg & Lee, 2017). By repeating this procedure with various variable values, a picture of the model's behavior can be illustrated. In numerous practical domains, ME results have been observed to be direct and intuitive.

Cameron and Trivedi (2005) define the ME of dummy and continuous variables with the following equation:

$$ME(X_k) = Pr(Y = 1 | X, X_k = 1) - Pr(Y = 1 | X, X_k = 0). \quad (1)$$

The equation for the ME of a continuous variable is defined in Equation (2):

$$ME(X_k) = \lim_{\Delta \rightarrow 0} \left[ \frac{\{Pr(Y = 1 | X, X_k + \Delta) - Pr(Y = 1 | X, X_k)\}}{\Delta} \right]. \quad (2)$$

Second, a rigorous discussion regarding the reliability of the estimated ME results is required; however, it is difficult for the estimated ME results of an ML-based model to be consistent. This is due to the non-linear or non-parametric characteristics typically inherent in an ML model, which causes high variance in the dependent variable values in ME depending on the location of the independent variable being estimated. In addition, if the dependent variable is discontinuous, it may not respond to Delta ( $\Delta$ ), which denotes infinitesimal changes in the independent variable. As a result, a *t*-test is conducted to explain the statistical significance of the changes observed in ME to check its reliability.

Since ML models take a non-linear form, infinitesimal changes in the independent variable may have no effect on the dependent variable (Bracke et al., 2019). To prevent such an occurrence, we identify the critical value at which there is a statistically significant difference in the default risk by applying various ranges of change, that is  $\Delta$ , to the independent variable. For continuous independent

variables, ME was estimated through 1% stepwise changes from 0 to 100% of the independent variable's standard error (SE) for  $\Delta$ . The procedure and code for the ME algorithm are both included in Algorithm 1 in the Appendix. Accordingly, a *t*-test was utilized to identify the critical value at which the null hypothesis of ME can be rejected at the 10% and 1% significance levels. Since it is unrealistic to apply  $\Delta$ , which represents infinitesimal changes, to the SE for dummy variables, we do not conduct the ME analysis for dummy variables.

By estimating ME, we identify the extent and direction of how the individual independent variables affect the model's dependent variable; further, it is possible to comprehend the range of change in the independent variable that has a statistically significant effect on the change in default risk probability.

### 3. Empirical results

#### 3.1. Dataset and variables

This study utilizes Freddie Mac's U.S. Single-Family Loan Dataset. Here, default is operationally defined as a loan that is at least 90 days delinquent (90 DPD). Unlike the competing risks model approach undertaken in prior econometric analyses, the ML approach is generally applied to a single default risk model (without prepayment risk analysis). The mortgages are homogeneous 30-year fixed-rate mortgages, such that there exists a borrower with 1 unit of property and a purchase-money mortgage. Table 1 exhibits a left censored dataset from 2011 to a right censored one in September 2013. The number of loans originated by year and number of yearly defaults during the observation period are provided in columns (A) and (B), respectively. There is a total of 1.102 3-year loan defaults, which is 6.93% of all observations.

The definitions of the independent variables and how they are calculated are provided in Table A.1 in the Appendix; the descriptive statistics of each variable are found in Table A.2 in the Appendix.

Table 1. Data structure and yearly defaults (source: authors' compilation from Freddie Mac's database)

Origination year (A)	Yearly origination	Default event year (B)			3-year default total
		2011	2012	2013	
2005	2.626	98	65	34	197
2006	2.677	135	77	45	257
2007	3.187	163	101	49	313
2008	3.441	132	87	42	261
2009	3.351	28	25	16	69
2010	622	2	1	2	5
Total	15.904	558	356	188	1.102



### 3.2. Logit model

To investigate the differences between the two main approaches and the possibility of mutual complementary support (Bücker et al., 2020; Chernozhukov et al., 2018; Mullainathan & Spiess, 2017; Rudin, 2019), we estimate a logit regression model (see Table A.3 in the Appendix). The unemployment rate, current loan to value (LTV), original debt to income (DTI), number of times thirty days delinquent in the last 12 months (30 DPD) and number of times sixty days delinquent (60 DPD) displayed positive (+) effects on default, while the origination 2008 and 2009 dummy, original LTV, credit score, power of sale states dummy, and sand states dummy displayed negative (-) relationships. However, the relationships of the mortgage age-related variables, house price appreciation rate, owner occupied, first-time home buyer, piggyback, and equity ratio-related variables were not found to be statistically significant.

While the model estimation results had no significant differences when compared to prior studies, differences were found in the case of original LTV. In existing literature, Foote et al. (2008) state that a higher original LTV (+) and interest rate (+) tend to raise the default probability. Moreover, Bracke et al. (2019) find that current LTV and original LTV have positive (+) effects on default risk in their logit model. However, the results of the original LTV in our model, which are contrary to common knowledge, may be explained in terms of endogeneity bias (Bhardwaj & Sengupta, 2010).

Table 2 presents the results of ME based on our logit regression model. An additional amount of  $\Delta$  disruption was induced in the current value of each independent variable to calculate the changes observed in the dependent variable. All the ME analyses were conducted using Python software with the help of the scikit-learn ML toolkit

and other ML libraries run on Ubuntu Linux and Amazon Web Services with 4 CPUs and 16 GB memory.

Further, a *t*-test was conducted after computing the mean of the ME vector and the SE of each continuous independent variable. To compare models, the size of the  $\Delta$  with respect to the SE of a variable was increased by 1% increments from 0% to 100% until ME (change in default probability) was statistically significant either at the 1% or 10% significance level. Columns (a) and (b) of Table 2 exhibit the ME values calculated using the critical values that reject the null hypothesis at the 1% statistical significance level. Columns (c) and (d) present the size of  $\Delta$  that rejects the null hypothesis of ME at the 1% or 10% confidence level. Here, we need to explain the economic implications of the results of the ME analysis in terms of mortgage age, house price appreciation rate, and equity to debt ratio, which are not statistically significant in the logit model estimation. Since the ME values of insignificant variables do not have any economic implications, we only add them for the purpose of comparing them with other ML models.

Table 2 shows that a disruption of 40% in the SE of the original DTI variable results in an ME of 0.1037, which is found to be statistically significant at the 1% confidence level. For the current LTV variable, a disruption size of 4% implies that the null hypothesis of ME is rejected at the 10% significance level. Therefore, compared to the original DTI variable, the ME of the current LTV variable is more sensitive to a small disruption (4% of the standard deviation of a variable). This means that even if the variables possess the same level of statistical significance, their ME sensitivities can differ widely. Overall, since the ME of the logit model is calculated through a logit regression, the direction of all variables remains the same but the scales of their responses do not appear to be uniform.

Table 2. ME of logit regression

Variable#	ME (Based on a confidence level of 1%)		Critical value as a Ratio of $\Delta$ to SE (in %)	
	Mean of ME* (a)	SE of ME (b)	Confidence level of 10% (c)	Confidence level of 1% (d)
Unemployment rate###	0.4701***	0.2384	14	22
Mortgage age	-0.3675***	0.1891	7	10
Square of mortgage age	-0.2069***	0.1065	10	16
House price appreciation rate	0.0955***	0.0484	38	59
Current LTV###	1.1218***	0.5683	4	7
Original DTI##	0.1037***	0.0526	25	40
Original LTV###	-0.2979***	0.1532	11	17
Credit score###	-0.6016***	0.3095	6	8
Equity to debt ratio	-2.027***	1.225	1	1
Square of equity to debt ratio	2.558***	1.108	1	1

Note: # and \* indicate the statistical significance level in the logit model estimation and the ME of the logit model, respectively.

### 3.3. ANN model

In complex non-linear ANN models, inducing separate changes in an individual variable has a limited effect on the dependent variable (Bracke et al., 2019).

Table 3 displays the results of calculating ME for an ANN model. ME was calculated using the same process as that used for the logit regression model. The ANN model applied in this study is a relatively simple network with a single hidden layer and an equal number of hidden nodes and independent variables. Despite its simplicity, the ANN model shown required relatively large changes in independent variables for its ME to be statistically significant.

We discuss the estimation results using an economic intuitive perspective. The credit score variable has played an important role in most standard econometric models (Archer & Smith, 2013; Bhardwaj & Sengupta, 2010; Elul et al., 2010), although it has been deemed unimportant in the case of ML models (Bracke et al., 2019). However, it is significantly important in Sirignano et al. (2016) and in our ANN model, credit score is found to be the 5<sup>th</sup> most important variable.

The original DTI does not play an important role in the ANN model estimation, although it is statistically significant in the ME analysis. We conjecture that the lack of importance of the economically important original DTI variable in the ML model might be the result of an effect that is similar to that of multicollinearity in standard econometrics; that is, the original DTI variable might be explained by other independent variables in a non-linear manner in the ANN model estimation process. This is certainly an advantage of the ME analysis, which provides ML modelers with both the directional signs of individual variable influences and the results of tests of statistical significance to consider further hand-tuning of the model.

In contrast to the ANN model, the ME analyses find that the coefficients from the logit model are more sensitive to small changes in the mean of the absolute values of ME. This is significant from the viewpoint of policy interventions, such as in the case of a predictive policing system (Datta et al., 2016). Since human regulators have limited capabilities for analyzing the influence of simultaneous changes in multiple variables and manipulating multiple variables jointly and marginally, it is more

Table 3. ME of ANN model

Variable#	ME (Based on a confidence level of 1%)		Critical value as a Ratio of $\Delta$ to SE (in %)	
	Mean of ME	SE of ME	Confidence level of 10%	Confidence level of 1%
Unemployment rate###	0.1148***	0.1677	26	40
Mortgage age	-0.0773***	0.1502	13	20
Square of mortgage age	-0.1188***	0.2019	8	12
House price appreciation rate	0.0144	0.1660	Not significant	Not significant
Current LTV###	0.4862***	0.6058	4	7
Original DTI##	0.0314***	0.0642	37	58
Original LTV###	-0.0502***	0.0687	29	45
Credit score###	-0.2013***	0.2829	7	11
Equity to debt ratio	-0.4798***	0.5867	56	87
Square of equity to debt ratio	-0.0823	0.0984	Not significant	Not significant

Note: # and \* indicate statistical significance levels of the logit model estimation and ME of the ANN model, respectively.

Table 4. Comparison of ME between logit regression and ANN model

Variable#	Logit model	ANN model	Sign	Significance level
Unemployment rate###	0.4701***	0.1148***		
Mortgage age	-0.3675***	-0.0773***		
Square of mortgage age	-0.2069***	-0.1188***		
House price appreciation rate	0.0955***	0.0144		Different
Current LTV###	1.1218***	0.4862***		
Original DTI##	0.1037***	0.0314***		
Original LTV###	-0.2979***	-0.0502***		
Credit score###	-0.6016***	-0.2013***		
Equity to debt ratio	-2,027**	-0.4798***		
Square of equity to debt ratio	2.558***	-0.0823	Opposite	Different

Note: # and \* indicate the statistical significance level in the logit model estimation and MEs of both the logit and ANN models, respectively.

practical to use a standard econometrics approach such as an ME analysis. Even though ML models consider the presence of many correlated input variables, the fact that the ML approach underestimates the influence of an individual input variable can result in serious problems.

Table 4 presents a comparison between the ME of the logit regression and ANN models. In contrast to the logit model, the house price appreciation rate and square of equity to debt ratio variables in the ANN model fail to achieve statistical significance level in the ME analysis. Given that the house price appreciation rate variable is the 10<sup>th</sup> most important variable under the ANN model, the results do not appear to be plausible. Therefore, while the two models are quite similar, the results of the logit model are slightly more consistent than those of the ANN model estimation.

By perturbing two variables simultaneously, we measured the combined ME on default probability. While the experiment can be conducted by grouping all variables, considering two variables at a time, this study focuses on the original LTV (ltv) and original DTI (dti), which are representative variables of the double trigger property (equity and cash-flow considerations). During this procedure, increasing the original LTV (ltv) and original DTI (dti) simultaneously was shown to increase default probability; therefore, both LTV and DTI were increased in the positive direction to identify the area where the default probability is found to exhibit a statistically significant change. We expect that in comparison to changing one variable, changing two variables simultaneously will result in a more sensitive response in the dependent variable, such that the two variables should display a relationship of substitution.

In contrast to our expectation, Panel A of Figure 1 shows complementary relationships. This is simply because the original LTV in the ANN model has a negative ME. Similarly, Sirignano et al. (2016) conduct analyses by pairing the original interest rate, interest rate differentials, original loan term, FICO score, loan balance, and past de-

linquency behavior, which were deemed to have high integrated effects. In their study, the interaction of the original interest rate and FICO score are explained with a contour plot in the presence of several regimes in the FICO score, such that for very high FICO scores, borrowers rarely become delinquent irrespective of how high the original interest rate is. In contrast, the study states that for low FICO scores, there exists a non-linear relationship with the interest rate, such that the likelihood of delinquency increases as the interest rate increases.

With the help of PDP, changes were observed in the dependent variable as a result of changes in the values of each independent variable. While the PDP of most variables displayed the same directions as that of ME, the PDP of the house price appreciation rate had a quadratic functional form, as shown in Figure 2. This may suggest that an additional sub-group analysis for house price appreciation rate values is necessary, following Athey and Imbens (2016), Chernozhukov et al. (2018), and Torrent et al. (2020). If these two sub-groups pass the Chow test, then we should follow this direction of sub-group analysis.

Changes in default probability resulting from changes in double trigger variables, such as the original LTV and original DTI, were estimated using the PDP technique. The results from the PDP analysis are similar to those of the ME analysis. If we increase original DTI (dti) and decrease original LTV (ltv) simultaneously, the default probability increases (Figure 3).

Further, the SHAP value allows an evaluation of each variable's marginal contribution by considering all correlations and interactions between them. However, this increases the time consumed for estimation. Taking this issue into consideration, a random sampling of 1,000 observation values was conducted to calculate the SHAP value. It is not possible to go through all the possible combinations of input variables (Štrumbelj & Kononenko, 2011).

The results of the SHAP value in the ANN model show that the current LTV (cu\_ltv) variable is of primary importance, followed by credit score, mortgage age, origination

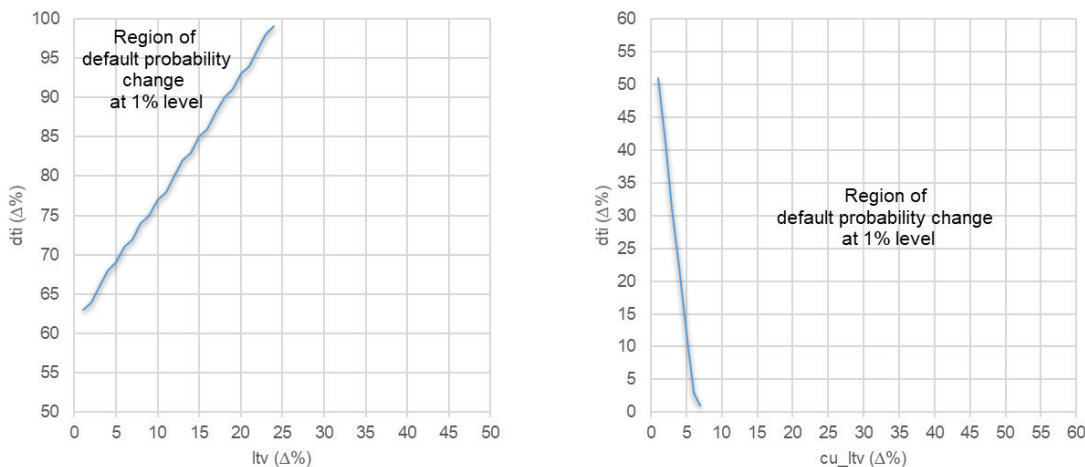


Figure 1. Change in ME resulting from a change in the original DTI (dti), and either original LTV (ltv) or current LTV (cu\_ltv) in the ANN model at the 1% significance level

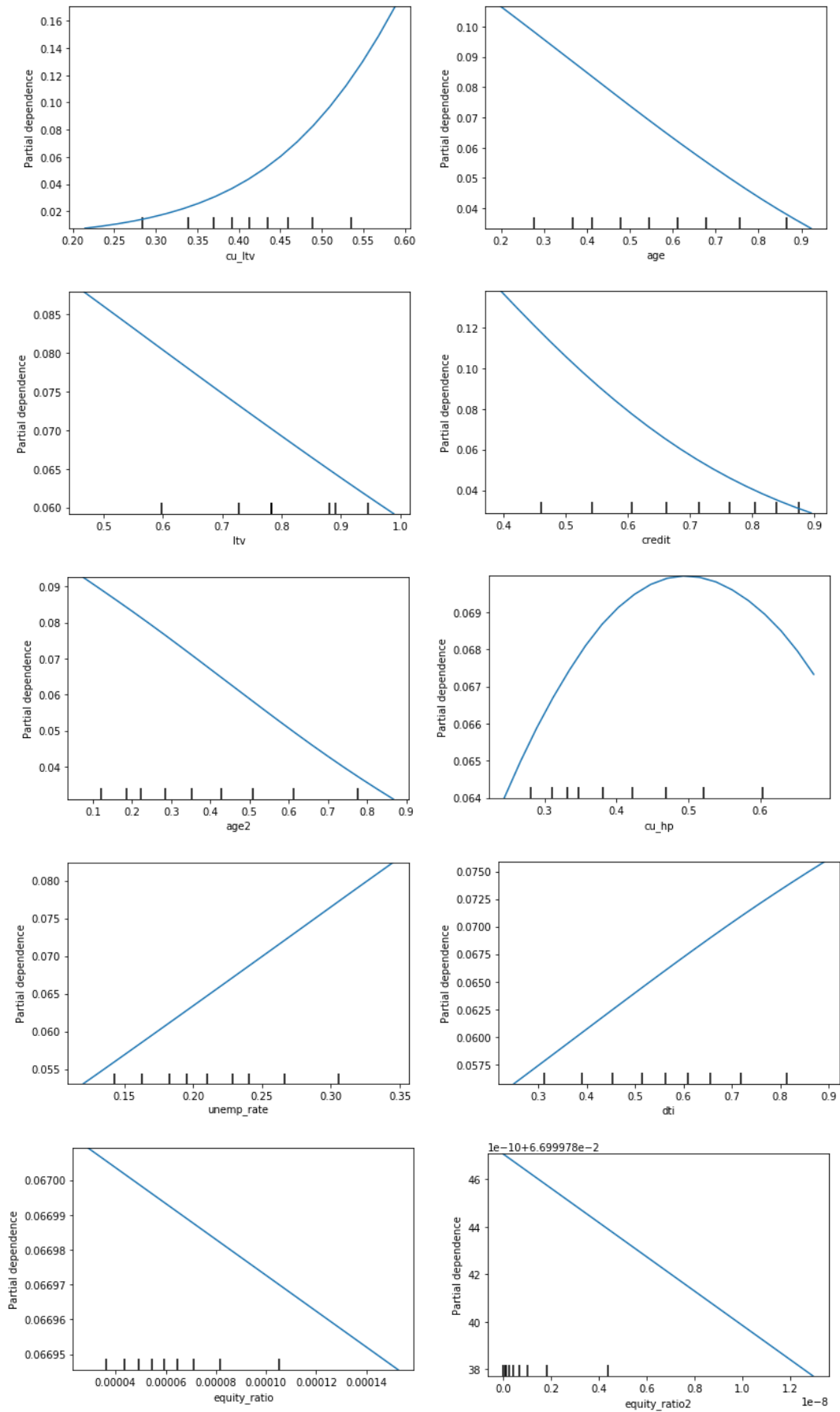


Figure 2. PDP of the ANN model



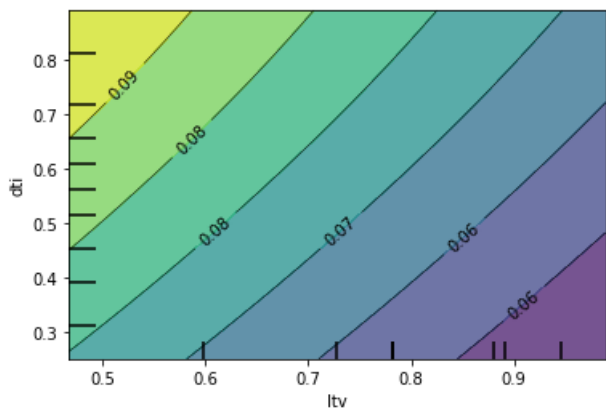


Figure 3. PDP resulting from changes in original LTV (ltv) and original DTI (dti) in the ANN model

2009, square of mortgage age, origination 2008, 30 DPD, and unemployment rate, in declining order of importance (Figure 4). Additionally, an analysis of the top 10 important variables provides the same results as those found in the logit estimation analysis. Therefore, these top 10 ANN model variables coincide with the logit model variables with statistical significance at the 1% level, except in the case of mortgage age and square of mortgage age. However, 60 DPD, with significance at the 1% level in the logit analysis, is weakly important in the SHAP value analysis of the ANN model. Furthermore, house price appreciation rate and equity to debt ratio, which were not statistically significant in the logit analysis, exhibited no or minimal effect in the SHAP value analysis of the ANN model.

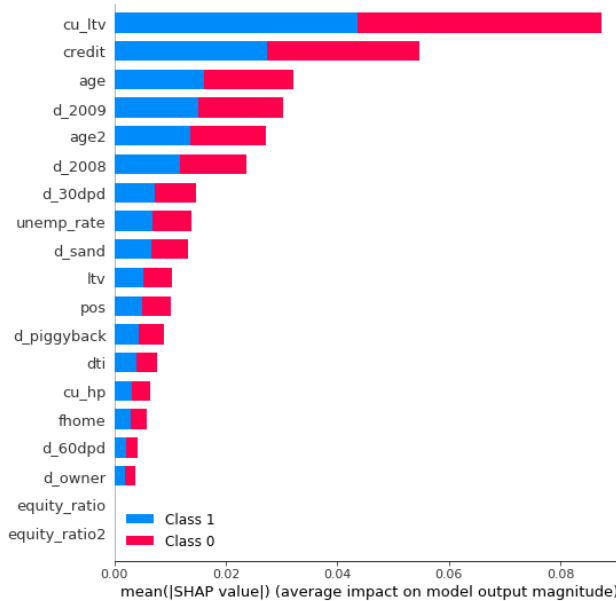
Bracke et al. (2019) utilize a logit model to calculate unary and SHAP values and find similar characteristics to those of their logit model. The SHAP value results from their study, in decreasing order of importance, indicate that year, current rate type, current LTV, current interest, outstanding balance, gross income, and LTV affect default risk.

When we compare the ME and SHAP value analyses of the ANN model, the house price appreciation rate and squared equity to debt ratio, which were not statistically significant in the ME analysis, also had no or minimal effects in the SHAP value analysis of the ANN model. Compared to the SHAP analysis, our proposed ME analysis is much easier to implement and understand. The results of the ME analysis provide both the directional sign and statistical significance, in addition to the size of the impact of the SHAP value analysis.

### 3.4. RF model

Table 5 displays the results of calculating the ME for the RF model. The process of applying the ME was similar to that in the logit regression and ANN models. To achieve continuity of change in the probability value, 100 decision trees were generated for this analysis.

Similar to the ANN-based model, which requires a relatively significant change for the ME, the RF model also has similar characteristics while displaying a degree of



Note: Class 1 indicates the default classification.

Figure 4. SHAP value results for the ANN model

sensitive response. We conjecture that it is a general property of an ML model, which allows for correlated changes among input variables, as compared to the logit model, which assumes the input variables are independent.

Table 6 compares the ME from the logit regression and RF models. While the original LTV had a negative ME in both the logit regression and ANN model, it the ME in RF was positive. Logically, it should positively affect defaults; therefore, the RF model is deemed to be more appropriate in economic theory.

We need to be cautious while considering that we have mutually contradictory ME results between models. Similar findings from a comparison among competing models were reported in Bracke et al. (2019). While year, current rate type, current LTV, current interest, and outstanding balance, in this decreasing order, were considered important in the logit model estimation, the variables in the Gradient Tree Boosting model appear in the order of year, current interest, current LTV, gross income, and current rate type, indicating wide differences between the models. While it may be difficult to generalize these results in a broad application, we assert that they would provide insights when selecting the independent variable for every default model in the spirit of hand-curating, following Mullainathan and Spiess (2017).

In addition to these results, we estimated how perturbing two variables simultaneously affects the default probability. The original LTV (ltv), current LTV (cu\_ltv), and original DTI (dti) were selected. The results between original DTI and current LTV are quite similar to those found in the ANN model, while the results of the original DTI and original LTV are quite different (Figure 5). This is due to the difference in the sign of the original LTV variables in the two ML models.

Table 5. ME of RF

Variable#	ME (Based on confidence level of 1%)		Critical value as a Ratio of $\Delta$ to SE (in %)	
	Mean	SE	10% confidence level	1% confidence level
Unemployment rate###	0.181***	1.4948	21%	37%
Mortgage age	0.0047	0.2822	Not significant	Not significant
Square of mortgage age	0.0075	0.2543	Not significant	Not significant
House price appreciation rate	0.0483***	0.5031	54%	74%
Current LTV###	0.9441***	3.8948	4%	5%
Original DTI##	0.0548***	0.3355	38%	55%
Original LTV###	0.0359***	0.3005	62%	92%
Credit score###	-0.1128***	1.0533	16%	27%
Equity to debt ratio	230.3042***	1084.8710	1%	1%
Square of equity to debt ratio	-0.1472	0.3906	Not significant	Not significant

Note: # and \* indicate statistical significance level for the logit model estimation and ME of the RF model, respectively.

Table 6. Comparison of ME between logit regression and RF models

Variable#	Logit model	RF model	Sign	Significance level
Unemployment rate###	0.4701***	0.181***		
Mortgage age	-0.3675***	0.0047	Opposite	Different
Square of mortgage age	-0.2069***	0.0075	Opposite	Different
House price appreciation rate	0.0955***	0.0483***		
Current LTV###	1.1218***	0.9441***		
Original DTI##	0.1037***	0.0548***		
Original LTV###	-0.2979***	0.0359***	Opposite	
Credit score###	-0.6016***	-0.1128***		
Equity to debt ratio	-2027.093***	230.3042***	Opposite	
Square of equity to debt ratio	2558.5336***	-0.1472	Opposite	Different

Note: # and \* indicate statistical significance level for the logit model estimation and MEs of both logit and RF models, respectively.

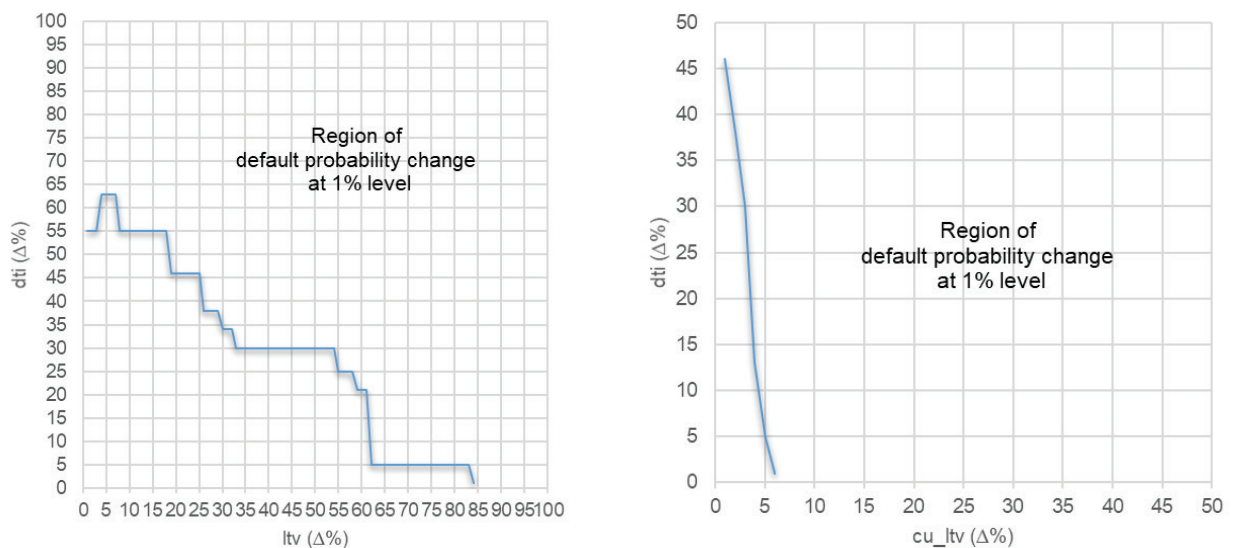


Figure 5. Changes in ME resulting from changes in original DTI (dti), and either original LTV (ltv) or current LTV (cu\_ltv) in the RF model at the 1% significance level

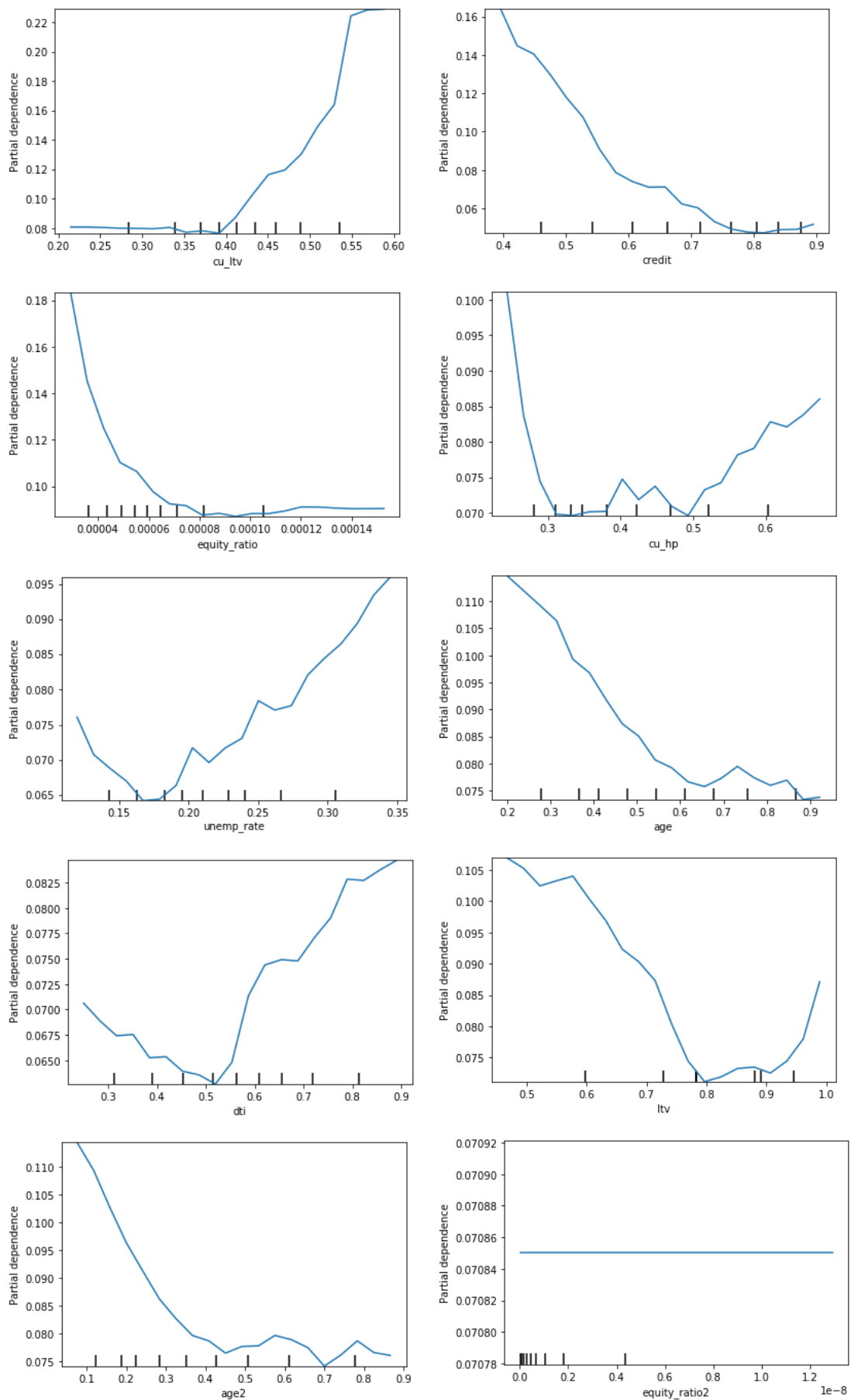


Figure 6. PDP of the RF model

Unlike the ANN model, the RF technique displays the characteristics of a discrete model based on decision trees, resulting in a complex, non-linear, quadratic PDP. Unlike the smoothly curved or linear form of ANN, drastically different forms were observed. Both positive and negative directions, depending on the position of the variable, were confirmed through PDP, as shown in Figure 6. Further, the original LTV displayed a negative direction in the logit and ANN analyses, while the direction is a quadratic functional form in the case of the RF model.

Although the variable had a positive direction in ME, both negative and positive directions can be confirmed via PDP. This signifies that under various circumstances, this variable can have different effects on the dependent variable. If we pass the Chow test after sub-grouping the sample data based on the PDP analysis, additional sub-group analysis will certainly improve the quality of standard econometric approaches.

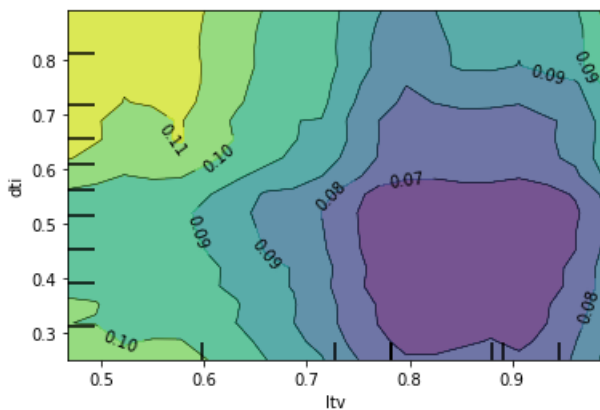
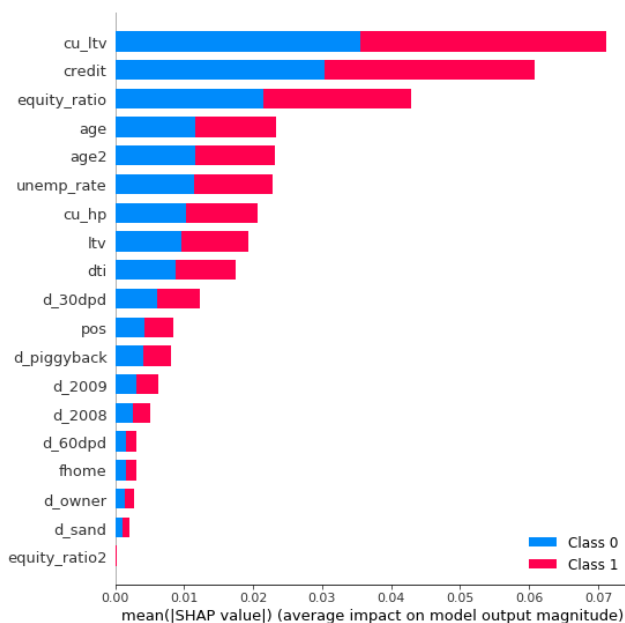


Figure 7. PDP for changes in original LTV (ltv) and original DTI (dti) in the RF model



Note: Class 1 means default classification.

Figure 8. Results of the SHAP value in the RF model

The changes in default probability resulting from changes in the double trigger variables (original LTV (ltv) and original DTI (dti)) were visualized in Figure 7. The PDP that responds to changes in both original DTI and original LTV showed a squiggly elliptical contour. This result differs from that of the ANN model and the RF model’s ME analysis.

Figure 8 displays the results of applying the SHAP value to the RF model. The top 10 most important variables from the SHAP value analysis indicate that the variables are also statistically significant in the RF model’s ME within a 1% level, except for two mortgage age-related variables.

Further, similar to our previous argument in the ANN model, our proposed ME analysis is much easier to implement and understand.

### 3.5. Model explainability analysis through the addition of polynomial variables

To analyze the changes in the independent variables within the black box of the ML models, the ML models were estimated by squaring and cubing all continuous variables (see Table A.1 in the Appendix). Further, the Mann-Whitney Test was conducted to confirm whether the artificial addition of polynomial variables (hereinafter referred to as the “polynomial models”) led to significantly different predictions compared to those of the original model.

Table 7. Mann-Whitney test results between the original and polynomial models

	Logit model	ANN model	RF model
by Classification*	0.062*	0.126	0.500
by Probability	0.259	0.000***	0.335

Note: \*We use the 50% rule: 1 if the probability is greater than or equal to 50%.

As Table 7 shows, based on default probability, of the polynomial models, only the ANN model differed from the original model with a significance level of 1%. In contrast, based on the classification results of default, the null hypothesis that there are no differences between the RF models was accepted. As a result, the argument that the ML model considers all possible function transformations is weakly rejected, which differs from Sirignano et al. (2016). Hence, modelers must continue to make diligent efforts in the selection of input variables in the spirit of hand-curating, following Mullainathan and Spiess (2017).

### Conclusions

This study employs both standard econometrics and ML techniques to estimate a residential mortgage loan default risk model and introduces economic explainability measures for its interpretation. We estimated the models’ explainability measures to identify the causal influences of individual independent variables through ME, PDP, and SHAP value analyses. Additionally, the significance levels



of the ME were utilized to conduct statistical tests on the ME of the respective variables on default probability.

Existing empirical studies on residential mortgage default have mainly explained default factors through an econometric approach. In contrast, studies based on ML highlight the high suitability of ML models in the context of predictive power. However, this study simultaneously employs both the explainability measures of ML techniques and the ME of standard econometric approaches to illustrate that our proposed methods can be utilized to develop a better mortgage default risk model. Specifically, for a standard econometric approach, analyses using ML approaches, such as PDP and SHAP value, can provide additional information regarding the need to sub-group analysis (Bücker et al., 2020; Chernozhukov et al., 2018; Mullainathan & Spiess, 2017; Rudin, 2019). For ML models, the ME analysis provides ML with a quasi-statistical inference technique with sign and statistical significance.

The results from comparing the logit regression, ANN, and RF models based on the evaluation of explainability show that while the PDP, ME, and SHAP values of logit regression and ANN display similar characteristics, the characteristics of the RF are different. This is due to the presence of structural characteristics in the RF model estimation: because RF employs a greedy heuristic approach, the difference between models do not depend on the number of trees.

This study enabled us to interpret the ML-based mortgage default risk model, which possesses the characteristics of a black box structure, in terms of ME, which has its foundations in economic theory. We expect our findings to play a practical guiding role when mortgage lending institutions utilize ML techniques for risk management of mortgage loans. From the perspective of policy formulation, our findings can be utilized by government regulators who are responsible for implementing macroprudential policies to manage household debt in the real estate financial market.

However, several limitations are associated with this study. First, it only uses event-month data instead of panel data with monthly observations. Second, we do not use a competing risks model that simultaneously deals with prepayment risk. Third, we face a number of contradictory results, depending on the estimation models, ME, and marginal importance techniques. Considering that both the standard econometric and ML models use the same data and share the same research purpose in terms of explainability and accuracy, the direction of future studies will be developing a unified research framework. We believe that our investigation can serve as a stepping-stone for a unified approach that can sufficiently offset this study's limitations.

## Acknowledgements

This paper was supported by the Konkuk University in 2020. The methodology used in this paper is based on the doctoral dissertation of the Dong-sup KIM. The authors

would like to thank two anonymous referees for their helpful suggestions in improving this study. Seungwoo Shin would like to take this opportunity to thank Professor Michael Seiler for his valuable comments.

## Funding

This work was supported by the Konkuk University under Grant in 2020.

## References

- Archer, W. R., & Smith, B. C. (2013). Residential mortgage default: the roles of house price volatility, euphoria and the borrower's put option. *The Journal of Real Estate Finance and Economics*, 46(2), 355–378. <https://doi.org/10.1007/s11146-011-9335-y>
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., & Herrera, F. (2020). Explainable artificial intelligence (XAI): concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Athey, S., & Imbens, G. (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences*, 113(27), 7353–7360. <https://doi.org/10.1073/pnas.1510489113>
- Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: causality and policy evaluation. *Journal of Economic Perspectives*, 31(2), 3–32. <https://doi.org/10.1257/jep.31.2.3>
- Bhardwaj, G., & Sengupta, R. (2010, October). *Where's the smoking gun? A study of underwriting standards for US subprime mortgages* (Working Paper). Federal Reserve Bank of St. Louis. <https://doi.org/10.2139/ssrn.1361139>
- Bracke, P., Datta, A., Jung, C., & Sen, S. (2019). *Machine learning explainability in finance: an application to default risk analysis* (Bank of England Working Paper No. 816). <https://doi.org/10.2139/ssrn.3435104>
- Bücker, M., Szepannek, G., Gosiewska, A., & Biecek, P. (2020). Transparency, auditability and explainability of machine learning models in credit scoring. *arXiv preprint arXiv:2009.13384*.
- Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. (2021). Explainable machine learning in credit risk management. *Computational Economics*, 57, 203–216. <https://doi.org/10.1007/s10614-020-10042-0>
- Cameron, A. C., & Trivedi, P. K. (2005). *Microeconometrics: methods and applications*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511811241>
- Campbell, J. Y., & Cocco, J. F. (2015). A model of mortgage default. *The Journal of Finance*, 70(4), 1495–1554. <https://doi.org/10.1111/jofi.12252>
- Chen, J. (2018). Fair lending needs explainable models for responsible recommendation. *arXiv preprint arXiv:1809.04684*.
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), C1–C68. <https://doi.org/10.1111/ectj.12097>
- Chopra, A., & Bhilare, P. (2018). Application of ensemble models in credit scoring models. *Business Perspectives and Research*, 6(2), 129–141. <https://doi.org/10.1177/2278533718765531>

- Dahiya, S., Handa, S. S., & Singh, N. P. (2017). A feature selection enabled hybrid-bagging algorithm for credit risk evaluation. *Expert Systems*, 34(6), e12217. <https://doi.org/10.1111/exsy.12217>
- Datta, A., Sen, S., & Zick, Y. (2016, May). Algorithmic transparency via quantitative input influence: theory and experiments with learning systems. In 2016 *IEEE Symposium on Security and Privacy (SP)* (pp. 598–617). IEEE. <https://doi.org/10.1109/SP.2016.42>
- Elul, R., Souleles, N. S., Chomsisengphet, S., Glennon, D., & Hunt, R. (2010). What “triggers” mortgage default? *American Economic Review*, 100(2), 490–494. <https://doi.org/10.1257/aer.100.2.490>
- Fahner, G. (2018). Developing transparent credit risk scorecards more effectively: an explainable artificial intelligence approach. In *The 7th International Conference on Data Analytics* (pp. 7–14), Athens, Greece.
- Feldman, D., & Gross, S. (2005). Mortgage default: classification trees analysis. *The Journal of Real Estate Finance and Economics*, 30(4), 369–396. <https://doi.org/10.1007/s11146-005-7013-7>
- Fitzpatrick, T., & Mues, C. (2016). An empirical comparison of classification algorithms for mortgage default prediction: evidence from a distressed mortgage market. *European Journal of Operational Research*, 249(2), 427–439. <https://doi.org/10.1016/j.ejor.2015.09.014>
- Foote, C. L., Gerardi, K., & Willen, P. S. (2008). Negative equity and foreclosure: theory and evidence. *Journal of Urban Economics*, 64(2), 234–245. <https://doi.org/10.1016/j.jue.2008.07.006>
- Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2020, October 1). *Predictably unequal? The effects of machine learning on credit markets* (Working paper).
- Galindo, J., & Tamayo, P. (2000). Credit risk assessment using statistical and machine learning: basic methodology and risk modeling applications. *Computational Economics*, 15(1), 107–143. <https://doi.org/10.1023/A:1008699112516>
- Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics*, 24(1), 44–65. <https://doi.org/10.1080/10618600.2014.907095>
- Kabul, I. K. (2017, December 18). Interpretability is crucial for trusting AI and machine learning. *The SAS Data Science Blog*. <https://blogs.sas.com/content/subconsciousmusings/2017/12/18/interpretability-crucial-trusting-ai-machinelearning/>
- Koutanaei, F. N., Sajedi, H., & Khanbabaie, M. (2015). A hybrid data mining model of feature selection algorithms and ensemble learning classifiers for credit scoring. *Journal of Retailing and Consumer Services*, 27, 11–23. <https://doi.org/10.1016/j.jretconser.2015.07.003>
- Kvamme, H., Sellereite, N., Aas, K., & Sjursen, S. (2018). Predicting mortgage default using convolutional neural networks. *Expert Systems with Applications*, 102, 207–217. <https://doi.org/10.1016/j.eswa.2018.02.029>
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 30, 4765–4774.
- Mullainathan, S., & Spiess, J. (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2), 87–106. <https://doi.org/10.1257/jep.31.2.87>
- Nanni, L., & Lumini, A. (2009). An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring. *Expert Systems with Applications*, 36(2), 3028–3033. <https://doi.org/10.1016/j.eswa.2008.01.018>
- Pearl, J. (2019). The seven tools of causal inference, with reflections on machine learning. *Communications of the ACM*, 62(3), 54–60. <https://doi.org/10.1145/3241036>
- Pławiak, P., Abdar, M., & Acharya, U. R. (2019). Application of new deep genetic cascade ensemble of SVM classifiers to predict the Australian credit scoring. *Applied Soft Computing*, 84, 105740. <https://doi.org/10.1016/j.asoc.2019.105740>
- Preece, A., Harborne, D., Braines, D., Tomsett, R., & Chakraborty, S. (2018). Stakeholders in explainable AI. *arXiv preprint arXiv:1810.00184*.
- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5), 206–215. <https://doi.org/10.1038/s42256-019-0048-x>
- Sealand, J. C. (2018). *Short-term prediction of mortgage default using ensembled machine learning models* [Unpublished doctoral dissertation]. Slippery Rock University.
- Shead, S. (2020, August 21). How a computer algorithm caused a grading crisis in British schools. *CNBC*. <https://www.cnb.com/2020/08/21/computer-algorithm-caused-a-grading-crisis-in-british-schools.html>
- Sirignano, J., Sadhwani, A., & Giesecke, K. (2016). Deep learning for mortgage risk. *SSRN Electronic Journal*.
- Steinwart, I., Hush, D., & Scovel, C. (2009). Learning from dependent observations. *Journal of Multivariate Analysis*, 100(1), 175–194. <https://doi.org/10.1016/j.jmva.2008.04.001>
- Štrumbelj, E., & Kononenko, I. (2011, April). A general method for visualizing and explaining black-box regression models. In *International Conference on Adaptive and Natural Computing Algorithms* (pp. 21–30). Springer. [https://doi.org/10.1007/978-3-642-20267-4\\_3](https://doi.org/10.1007/978-3-642-20267-4_3)
- Torrent, N. L., Visani, G., & Bagli, E. (2020). PSD2 explainable AI model for credit scoring. *arXiv preprint arXiv:2011.10367*.
- Tsai, C. F., & Wu, J. W. (2008). Using neural network ensembles for bankruptcy prediction and credit scoring. *Expert Systems with Applications*, 34(4), 2639–2649. <https://doi.org/10.1016/j.eswa.2007.05.019>
- Wallace, N. (2005). Innovations in mortgage modeling: an introduction. *Real Estate Economics*, 33(4), 587–594. <https://doi.org/10.1111/j.1540-6229.2005.00131.x>
- Wang, G., Ma, J., Huang, L., & Xu, K. (2012). Two credit scoring models based on dual strategy ensemble trees. *Knowledge-Based Systems*, 26, 61–68. <https://doi.org/10.1016/j.knosys.2011.06.020>
- Wang, J., Wiens, J., & Lundberg, S. (2021, March). Shapley flow: a graph-based approach to interpreting model predictions. In *International Conference on Artificial Intelligence and Statistics* (pp. 721–729). PMLR.
- Zhao, Q., & Hastie, T. (2021). Causal interpretations of black-box models. *Journal of Business & Economic Statistics*, 39(1), 272–281. <https://doi.org/10.1080/07350015.2019.1624293>

## Appendix

Algorithm 1. Algorithm for ME calculation

Procedure:

for variable  $i$

if type of variable  $i$  is continuous:

Calculate standard deviation ( $x_i^{std}$ ) of variable  $i$  ( $x_i$ )

Set  $\Delta$  as  $x_i^{std} / 100$

for  $j$  in [1, 100]

Change value of variable  $i$  ( $x_i$ ) as  $x_i + (j \times \Delta)$  (Set as new dataset)

Measure probabilities of model for original and new dataset, respectively

Calculate marginal effect using the difference in probabilities

Identify change in probabilities for the two datasets (original & new) using a  $t$ -test

Code:

```
def getME(estimator=False):
```

```
    if estimator == False:
```

```
        print('Estimator is not defined')
```

```
        return 0
```

```
    else:
```

```
        print('Estimator : ', estimator)
```

```
    for i in inputList:
```

```
        inputDataTmp = inputData.copy()
```

```
        if varType[i] == 'continuous':
```

```
            for j in range(0,100):
```

```
                tmpProb = estimator.predict_proba(inputData)[:,1]
```

```
                inputDataTmp[i] = inputData[i] + float(j+1)*varStd[i]
```

```
                tmpProbP1 = estimator.predict_proba(inputDataTmp)[:,1]
```

```
                tStat = sp.stats.ttest_ind(tmpProbP1, tmpProb).statistic
```

```
                pVal = sp.stats.ttest_ind(tmpProbP1, tmpProb).pvalue
```

```
                if j+1 == 100:
```

```
                    print(i, ', ', j+1, ', continuous, varAve, ', round(np.mean(inputData[i]),5),
```

```
                          ', increment, ', round(float(j+1)*varStd[i],5),
```

```
                          ', ave, ', round(np.mean((tmpProbP1-tmpProb)/(float(j+1.)*varStd[i])),5),
```

```
                          ', stdev, ', round(np.std((tmpProbP1-tmpProb)/(float(j+1.)*varStd[i])),5),
```

```
                          ', t-stat, ', round(tStat,5),
```

```
                          ', p-value, ', round(pVal,5))
```

```
                if pVal <= 0.1:
```

```
                    print(i, ', ', j+1, ', continuous, varAve, ', round(np.mean(inputDataTmp[i]),5),
```

```
                          ', increment, ', round(float(j+1)*varStd[i],5),
```

```
                          ', ave, ', round(np.mean((tmpProbP1-tmpProb)/(float(j+1.)*varStd[i])),5),
```

```
                          ', stdev, ', round(np.std((tmpProbP1-tmpProb)/(float(j+1.)*varStd[i])),5),
```

```
                          ', t-stat, ', round(tStat,5),
```

```
                          ', p-value, ', round(pVal,5))
```

```
                    if pVal <= 0.01:
```

```
                        break
```

```
    return 1
```

Table A.1. Variable names and descriptions

Variable	Symbol	Definition
Sixty days delinquent	d_60dpd	Equals 1 if there exists at least one 60 DPD, excluding the previous two months while considering the last 12 months; 0 otherwise
Thirty days delinquent	d_30DPD	Equals 1 if there exists at least one 30 DPD and only 30 DPD, excluding the previous month while considering the last 12 months; 0 otherwise. It excludes cases of 60 DPD
Unemployment rate	unemp_rate	MSA level rates provided by the United States Bureau of Labor Statistics
Mortgage age	age	Number of months since origination
Square of mortgage age	age2	Square of mortgage age
House price appreciation rate	cu_hp	Monthly house price appreciation rate of return
Current loan to value	cu_ltv	Current loan to value
Origination 2008	d_2008	Origination year dummy
Origination 2009	d_2009	Origination year dummy
Owner occupied	D_owner	Owner occupied state dummy
Original debt to income	dti	Debt to income ratio at origination
Original loan to value	ltv	Loan to value ratio at origination
Credit score	credit	Credit score at origination (provided by Freddie Mac)
First time home buyer	fhome	Equals 1 if when a buyer buys a house for the first time, and 0 otherwise
Power of sale states	pos	States with the power of sale foreclosure provision
Piggyback dummy	d_piggyback	Piggyback mortgage dummy
Sand states	d_sand	Arizona, California, Florida, and Nevada
Equity to debt ratio	equity_ratio	Equity to unpaid mortgage balance
Square of equity to debt ratio	equity_ratio2	Squared equity to debt ratio

Table A.2. Descriptive statistics

Variable	Remarks	Minimum	Maximum	Mean	Standard deviation
Sixty days delinquent	Dummy	0.00	1.00	0.01	0.08
Thirty days delinquent	Dummy	0.00	1.00	0.04	0.19
Unemployment rate	%	2.10	29.23	8.05	2.01
Mortgage age	Month	13.00	103.00	62.87	19.55
Square of mortgage age	Month	169.00	10,609.00	4,335.23	2,523.91
House price appreciation rate	Real number	-0.02	0.04	0.002	0.01
Current loan to value	Real number	0.00	1.93	0.79	0.22
Origination 2008	Dummy	0.00	1.00	0.22	0.41
Origination 2009	Dummy	0.00	1.00	0.21	0.41
Owner occupied	Dummy	0.00	1.00	0.87	0.33
Original debt to income	%	1.00	65.00	37.04	12.01
Original loan to value	%	8.00	100.00	80.20	13.73
Credit score	Real number	501.00	844.00	737.23	54.41
First time home buyer	Dummy	0.00	1.00	0.33	0.47
Power of sale states dummy	Dummy	0.00	1.00	0.56	0.50
Piggyback dummy	Dummy	0.00	1.00	0.13	0.34
Sand states	Dummy	0.00	1.00	0.15	0.36
Equity to debt ratio	Real number	-0.48	12,393.03	1.72	98.88
Square of equity to debt ratio	Real number	0.00	153,587,167.79	9,779.27	1,217,877.52



Table A.3. Estimation results: logit model

Variable	Coefficient	Standard error
Sixty days delinquent	6.80***	0.30
Thirty days delinquent	0.74**	0.29
Unemployment rate	0.14***	0.04
Mortgage age	-0.03	0.02
Square of mortgage age	-0.00	0.00
House price appreciation rate	13.85	9.24
Current loan to value	3.47***	0.77
Origination 2008	-1.36***	0.22
Origination 2009	-2.16***	0.32
Owner occupied	0.22	0.25
Original debt to income	0.01**	0.01
Original loan to value	-0.03***	0.01
Credit score	-0.004***	0.00
First time home buyer	-0.16	0.16
Power of sale states dummy	-0.30**	0.15
Piggyback dummy	0.10	0.22
Sand states	-0.61**	0.27
Equity to debt ratio	-0.81	0.57
Square of equity to debt ratio	0.00	0.00
Constant	-1.39	1.82

Note: \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.