Can Machine Learning Predict Dynamic Capital Structure?
- Evidence from Korea

Jin-Hyung Cho*

College of Economics and Finance, Hanyang University

Jeong-Hwan Lee**

College of Economics and Finance, Hanyang University

Bong-Jun Kim

College of Economics and Finance, Hanyang University

Won-Eung Lee

College of Economics and Finance, Hanyang University

Abstract

Yes, it can. Employing a variety of Machine Learning (ML) algorithms, we predict optimal capital structure of listed firms in Korea, comparing the performance of linear and machine learning models - namely, Multi-regression, LASSO, Random Forest (RF) and Gradient Boosting Regression (GBM). For analysis, we set the training and test set as 2003-2014 and 2015-2019 respectively. We find that the predicting performance on firm leverage, as measured in out- R^2_{0S} and MSE (Mean Square Error) for RF and GBM is much effective than that of LM and LASSO. In particular, the variables with high predictive power are the Market-to-Book ratio, NetPay, Z-score, Profit, and so on. Finally, after estimating the speed of adjustment (SOA) to the optimal capital structure, using the model of Amini et al. (2021), we confirm that RF and GBM are more predictive than LM and LASSO. Lastly, when chaebols, unique form of conglomerate in Korea, compared with non-chaebols, we find that the leverage adjustment speed of latter is much faster than that of the former especially in machine learning models which is due to the debt-dependent characteristic of *chaebols.*

Keyword: Machine Learning, Korea, Leverage, Leverage Adjustment Speed, Optimal Capital Structure

JEL Classification:

* Corresponding Author, Phd Candidate, College of Economics and Finance, Hanyang University, enish27@hanyang.ac.kr

** First Author, Associate Professor, College of Economics and Finance, Hanyang University, jeonglee@hanyang.ac.kr

1. Introduction

The determinants of capital structure have been a subject of discussion for last decades (Aybar-Aria, 2012; Baker et al., 2002; Fama et al., 2002; Flannery et al., 2006; Frank et al., 2004; Frank et al., 2009; Morellec, 2004; Myers, 1984). Despite extensive debate and research over last decades, the factors causing variations in capital structure remain a puzzle in developing countries including Korea (Choi, 2015; Kim et al., 2007; Lee et al, 2000; Son et al, 2007; Yoon, 2016). In fact, the capital structure and leverage of firms in Korea has been a significant matter of debate since the 1997 Asian Financial Crisis. The excessive investment, encouraged by financial institutions' lending led to low profitability and high leverage of these firms (Lee et al., 2000), which became a structural problem that caused the entire economy vulnerable. Among the firms, *chaebols*, unique form of conglomerates with historical ties to Korean government, are known to have extremely high leverages when compared to other firms. In this context, to better observe the speed of leverage adjustment of Korean firms since the 1997 Asian Financial Crisis, it is important to analyze Korean firms in two categories – *chaebols* and non-*chaebols* – to better observe dynamics in capital structure which would be unique to Korean business environment.

In particular, in contrast to previous researches, our study employs machine learning methodologies including Random Forests (RF), Gradient Boosting Machines (GBM) and Least Absolute Shrinkage and Select Operator (LASSO) models to address nonlinearities and interaction effects raised by previous researches in explaining the capital structure of Korean firms. We exploit a large sample which comprises 6,545 firm-year observations from 2003 to 2014, which split into 1,454 *chaebols* and 5,101 non-*chaebols*, and then use a training and cross-validation period from 2015 to 2019. Then, we assess out-of-sample R-squared (hereafter R^2_{0S})and outofsample mean squared forecast error (MSE) from 2015 to 2020 in order to assess the predictive power of our machine learning models relative to conventional linear models in measuring target leverages and evaluating their determinants. Subsequently, we estimate the speed of adjustment to target leverage and analyze whether machine learning models *better* predict observed firms' financing actions than traditional linear models.

The results of our analysis are as follows. First, our machine learning approach achieves substantial gains in estimating target leverage, R^2_{0S} statistics are 26.42% and 12.78% for OLS and LASSO over the testing period (2015-2019), compared to 55.29% and 52.85% for RF and GBM respectively. Thus, it seems clear that RF and GBM outperform linear models and lead to lower error every year in the testing period, especially for firm-fixed model. Moreover, machine learning models are found to significantly reduce MSE (Mean Square Error). For instance, RF and GBM lead to a 2.4% and 2.6% decline in MSE, compared to 14.9% and 12.66% in OLS and LASSO respectively.

Then, we turn to the assessment of the most important determinants of corporate leverage. In contrast to prior studies which employed in-sample tests to assess the importance of corporate leverage determinants, we conduct out-of-sample tests to exploit complex and high-dimensional patterns in leverage behavior and quantify the importance of its determinants. Our analysis proves that machine learning models capture information from a broader set of characteristics than previous studies. For instance, while RF show that firm-specific factors, such as Market-to-Book ratio (*Mktbk*), Net Payout (*Netpay*), Profitability (Profit), Bankruptcy Probability *(Z-score)* and Cash *(Cash)* are the primary determinants of market leverage, the rest of the models imply that macroeconomic factors, rather than firm factors are the determinants of market leverage.

Most importantly, we present evidence on how machine learning models achieve target leverage faster than traditional linear models. Here, our primary interest is not only the machine learning models' better performance of leverage adjustment speed, but also, to examine whether their estimate better conform to the trade-off theory, which suggests that firms' leverage adjustment speed completely conforms to the theory when it is close to 1, and 0 otherwise. In particular, we pay attention to whether the affiliation to *chaebols* play as a decisive role in firms' speed of leverage adjustment. Our result shows that the leverage adjustment speed of non-*chaebols* is much faster than that of *chaebols* in every model used including machine learning and traditional models; however, the leverage adjustment speed for non-*chaebols* under RF model is fastest at 48.27% which translates into 1.0515 half-life period, while the speed for *chaebols* under the same model is 30.87% which translates into 1.8777 half-life[1]. This could be due to debt-dependent characteristic of *chaebols* which is known to have high leverages when compared to non-*chaebols* (Lee et al, 2000).

The contribution of our research is as follows. First, to our knowledge, this study is first research to employ machine learning methodologies in analyzing target leverage for Korean firms. Second, we verified how machine learning models better estimate leverage adjustment speed and half-life year than traditional models, thereby better conforming to the trade-off theory. Lastly, we find the effect of heterogeneity between *chaebols* and non-*chaebols*, on *both* leverage adjustment speed and half-life year.

The rest of paper is organized as follows. Section 2 discusses previous researches and Section 3 presents data, methodology and predictive performance. Section 4 introduces machine learning and linear models used in our study. Section 5 presents the model for leverage adjustment speed and discusses the possible explanations for the results. Section 6 presents additional analysis for *chaebols* and non-*chaebols* and Section 7 concludes.

2. Literature Review

2.1. Debates on capital structure

The most prominent theories on capital structures are threefold: Trade-off theory, Pecking Order theory, and Market Timing theory. First, the trade-off theory suggests that there exists trade-off between benefits and costs that occurs from debt. In other words, firms decide their capital structure only taking into consideration the advantages and disadvantages of debt

[1] The detailed analysis will be presented in Section 7 which focus on the comparative analysis of *chaebols* and non-*chaebols*.

(Flannery et al., 2006; Frank et al., 2004; Graham et al., 2001). According to the theory, there exists target leverage which maximizes firm value and if this deviates, firms would try to adjust their leverage in effort to conformation to the target leverage. Second, the pecking order theory states that firms prefer to finance itself internally through retained earnings (Fama et al, 2002; Shyam-Sunder et al., 1999). However, if this source of financing is unavailable, they would finance through debt. And as a last resort, firms would choose to finance itself through the issuing of new equity. Lastly, market time theory states that in consideration of financing cost, firms use debt when the stock price is undervalued.

Among aforementioned theories, a large number of previous researches employ trade-off theory as a theoretical framework in analyzing target leverage (Fama et al., 2002; Flannery et al., 2006; Kim et al., 2007; Lemmon et al. 2008; Yoon et al, 2016). In general, it is widely believed that firms' leverage adjustment speed completely conforms to the theory when it is close to 1, and 0 otherwise. In fact, firms do consider their own target leverage for a variety of reasons. For instance, Graham et al. (2001) states that 81% of firms answered that they take target leverage into account in decision-making for loan. As well, after conducting survey, Bancel et al. (2004) conclude that 75% of European firms have their own target leverage. However, it seems that optimal target leverage implied by the trade-off theory is not easily achieved. As the volatility of stock prices frequently causes firms' market-to-debt ratio to deviate from its target, such deviations could cause firms not to immediately return to their target leverage by issuing or repurchasing securities (Eugene et al., 1977). For example, using OLS, Fama et al. (2002) finds that the adjustment towards target leverage is around 7-18% for a year. On the other hand, Flannery et al. (2007) and Lemmon et al. (2008) argue that firms tend to make a partial adjustment each year, moving around 30% of the way toward their target leverage.

2.2. Korean firms' capital structure

We choose the Korean firms to examine their capital structure for several reasons. First, Korean firms have had extremely high leverage, particular when compared to firms in other countries. As of the end of the 1997 Asian financial crisis. The total debt owed by Korean firms mounted up to 811 trillion won, which is equivalent to US$ 675 billion with a won/dollar exchange rate of 1,200, and was about 1.9 times as great as that year's GDP (Nam et al., 1999) The average debt/equity ratio of the 30 largest *chaebol*-affiliated firms exceeded 500% and some of them reached 3,000 percent (Lee et al., 2000). In the light of these considerations, we believe that there would exist an effect for heterogeneity for *chaebols* due to the history of leverage changes. Second, the estimated extents to which leverage adjustment speed are quite contrasting. Previous researches have employed a variety of econometric methodologies including firm-fixed panel model, OLS, FM, and GMM (Kim et al, 2007; Yoon et al, 2016; Lee et al, 2001). Using OLS and Fama et al. (1973) model, Yoon et al. (2016) find that the leverage adjustment speeds of Korean firms are 0.228 and 0.229 respectively. Also, Kim et al. (2007) note that the leverage adjustment speed is at 0.35 when firm-fixed model is employed. On the other hand, Kim et al. (2010) argues that leverage adjustment speed increases when the debt and equity financing market accessibility is fine; this result holds regardless of the use of book-to-market debt ratio as dependent variable. He further notes that firms can adjust leverage toward the target faster with the market accessibility of equity financing than that of debt financing.

Another branch of concerns are raised due to nonlinearity, or complex, high-order interaction effects (Strob et al, 2008; Altman et al, 2017) or even unbalanced data panel and endogeneity between independent variables are not properly taken into account seriously (Giovanni et al., 2005; Yoon et al, 2016). In this regard, we confirm the success of machine learning algorithms in predicting target leverage and financing actions in a highly nonlinear and discrete environment where firms with unusually high leverages are to be observed.

3. Methodology

We describe empirical models for predicting capital structure dynamics. The basic regression problem is to estimate a function $g(X_{i,t}) = E(y_{i,t+1} \mid X_{i,t})$ where

$$y_{i,t+1} = g(x_{i,t}) + \varepsilon_{i,t+1} \ldots (1)$$

is the $_i$th firm's target leverage ratio in year $_{t+1}$ and $\varepsilon_{i,t+1}$ is a random error component. Here, our goal is to estimate the function $g(\cdot)$ using machine learning and traditional regression methods. We expect that machine learning prediction functions capture nonlinear and discontinuous relations between y_i and other associated covariates. As well, we believe that these machine learning models may also explain complex interaction effects that could exist in the covariates.

3.1. Linear models

3.1.1 Multi-regression model

First, our baseline model is the following standard, multiple regression model of the form:

$$g(X_{i,t}\,;\,\beta) = X'_{i,t}\beta \ldots (2)$$

The regression parameters β estimated using OLS (Ordinary least squares) are defined by:

$$\hat{\beta}^{ols} = \text{argmin} \| \quad \|_2^2 \ldots (3)$$

where $\| a - b \|_2$ is the distance between the vectors a and b.

3.1.2. LASSO

Next, we present and utilize LASSO model to address drawbacks to the previous multiple regression: Overfitting and potential multi-collinearities. LASSO is known as a method for shrinking parameters related to insignificant covariates to zero (Tibshirani, 1990). Simply speaking, LASSO acts as both a shrinkage *and* a model selection tool. The parameter estimates in the LASSO model for a given value of λ are given by:

$$\hat{\beta}_{\lambda}^{\text{LASSO}} = \text{argmin}\| y - x\beta\|_2^2 + \lambda \| \beta \|_1 \text{ for some } \lambda > 0 \ldots (4)$$

3.2. Machine learning model

Machine learning algorithms are powerful for modeling nonlinear relationship among dependent and independent variables as well as capturing hidden complex interactions among them (Frank et al., 2009). The machine learning models we employ are random forest (Breiman, 2001) and gradient boosting machine (Friedman, 2001).

3.2.1. Random forests

Suppose that we split the feature space $\mathscr{X}$ into J unique nonoverlapping regions: R1, R2, …, Rj. The predicted value of y for any value within Rj is the average overall response values in Rj:

$$\hat{g}^{rf}(x) = \sum_{j=1}^{J} \bar{y}_j I_{\{x \in Rj\}} \ldots (5)$$

Where $I_{\{x \in Rj\}}$ is an indicator that equals one if x is in Rj and 0 otherwise. The predictions from growing a single tree are known to exhibit high variance. Thus, the method of bootstrap aggregation is employed to alleviate this potential problem (Altman et al., 2017; Breiman, 1996; Efron et al., 1994). The results in an ensemble of trees from which to make predictions. The bagged estimate at x is the average estimate over all trees:

$$\hat{g}^{bag}(x) = \frac{1}{B} \sum_{b=1}^{B} \hat{g}^{b}(x) \ldots (6)$$

where $\hat{g}^{bag}(x)$ is the estimator defined in Eq. (6) on the bth bootstrap sample.

3.2.2. Gradient boosting model

Compared with RF, gradient boosting regression trees sequentially grows trees by updating the data used to grow a tree after each tree is fit (Friedman, 2001). Starting by fitting a tree using the original data set, subsequent trees are grown using the fitted residuals, being updated after each fit. Then, the estimate of $\hat{g}_{gbm}(X)$ is the weighted sum of the individual estimates from each tree. Here, the weighting is controlled by a parameter λ that determines how fast the model "learn".

3.3. Model tuning and fitting

Machine learning models have one or more parameters not directly estimable from the data, often referred to as tuning parameters. A candidate set of parameters can be used a priori to assess their effectiveness. For each candidate parameter, the model is fit to a subset of the data and then assessed on its performance at predicting new observations in which the user knows the true response variables. To assess a parameter's effectiveness on model performance, we use subsampling from the original data set. We divide our data into training sets (2003-2014) and testing sets (2015-2019). Within the training set, we create 00 subsamples in the ML parlance, upon which we turn and validate each model's performance on a validation set.

One the appropriate set of tuning parameters is set for its respective model, we fit each candidate model to the entire training set from 2003 to 2014 using these tuning parameters. Then we use this model to make out-of-sample predictions in subsequent years. To be specific, the model would be updated in each subsequent year to make predictions on the following year. This process will be repeated through 2019.

4. Sample selection and summary statistics

4.1. Sample selection

Now we explain sample distribution and variable definition. Our primary sample includes Korean firms listed on KOSPI, which are covered by Dataguide between 2003 and 2019. We exclude firms affiliated industries in finance or insurance. Subsequently, the Table 1 below presents observation number for all firms, *chaebols* and non-*chaebols* by each year. In total there are 6,545, 1,454, 5,101 firm-level observations respectively.[2]

Table 1. Sample distribution for all firm, chaebols and non-chaebols

All firm				Chaebol				Non-chaebol			
Year	Number of Firm	Percent (%)	Cumulative Percent (%)	Year	Number of Firm	Percent (%)	Cumulative Percent (%)	Year	Number of Firm	Percent (%)	Cumulative Percent (%)
2003	318	4.86	4.86	2003	60	4.13	4.13	2003	258	5.06	5.06
2004	317	4.84	9.7	2004	65	4.47	8.60	2004	252	4.94	10.00

[2] The analysis in Section 4 will focus on the sample of all firms. The detailed analysis for chaebols and non-chaebols will be conducted in following Section 5.

2005	335	5.12	14.82	2005	69	4.75	13.35	2005	266	5.21	15.21
2006	343	5.24	20.06	2006	77	5.30	18.64	2006	266	5.21	20.43
2007	375	5.73	25.79	2007	84	5.78	24.42	2007	291	5.70	26.13
2010	407	6.22	32.01	2010	89	6.12	30.54	2010	319	6.25	32.39
2011	437	6.68	38.69	2011	103	7.08	37.62	2011	335	6.57	38.96
2012	448	6.84	45.53	2012	113	7.77	45.40	2012	336	6.59	45.54
2013	455	6.95	52.48	2013	117	8.05	53.44	2013	339	6.65	52.19
2014	465	7.1	59.59	2014	115	7.91	61.35	2014	351	6.88	59.07
2015	492	7.52	67.1	2015	112	7.70	69.05	2015	381	7.47	66.54
2016	507	7.75	74.85	2016	116	7.98	77.03	2016	392	7.68	74.22
2017	528	8.07	82.92	2017	101	6.95	83.98	2017	428	8.39	82.61
2018	554	8.46	91.38	2018	114	7.84	91.82	2018	441	8.65	91.26
2019	564	8.62	100	2019	119	8.18	100	2019	446	8.74	100
Total	6,545	100		Total	1454	100		Total	5101	100	

Subsequently, we present the main variables used for our prediction in Table 2. Following Amini et al. (2021), we use four dependent variables in measuring leverage, which are TDM, TDA, LDA and LDA. Subsequently, firm-value and firm-specific factors include Profitability(*Profit*), Firm Size(*Assets*), Mature firm(*Mature*), Market-to-Book ratio (*Mktbk*), Assets Growth(*ChgAsset*), Physical Investment(*Capex*), Assets Tangibility(*Tang*), Innovation Investment(*RD*), Uniqueness(*Unique*), Nonproduction cost(*SGA*), Top Tax Rate(*Taxrate*), Depreciation(*Depr*), Stock Variance(*StockVar*), Bankruptcy Probability(*Zscore*), Debt Rating(*Rating*), Stock Returns(*StockRet*), Market Returns(*CrspRet*), Industry Leverage(*Industlev*), Industry Growth(*Industrgr*), Logarithm of annual firm profits (*Macroprof*), Net Payout(*Netpay*). Macroeconomic factors include Term Spread (*Termsprd*), Inflation (*Inflation*) and Growth in GDP(*Macrogr*). Industrial dummies are *Size Dummies, Growth Dummies* and *High-Tech Dummy*. Lastly, MVE and MVA are calculated to make up the variables such as Leverage measured in MVA, TDM and LDM accordingly. The detailed definition for each variable is presented in following Table 2.

Table 2. Variable Definitions

Variables	Description
Market Value of Equity(MVE)	The stock's fiscal year close price times common shares outstanding
Market Value of Assets(MVA)	Debt in current liability plus long-term debt plus preferred stock liquidating value minus deferred tax + MVE
Leverage(TDM)	(Debt in current liabilities plus long-term debt)/MVA
Leverage(TDA)	(Debt in current liabilities plus long-term debt)/total assets
Leverage(LDM)	Long-term debt/MVA
Leverage(LDA)	Long-term debt/total assets
Profitability(Profit)	Operating income before depreciation/total assets
Firm Size(Assets)	The logarithm of total assets
Mature firm (Mature)	A dummy variable that equals one if the firms has been in the Dataguide DB for last 5 years and equals zero otherwise

Market-to-Book (Mktbk)	Market value of assets/total assets
Assets Growth (ChgAsset)	Change in the logarithm of total assets
Physical Investment(Capex)	Capital expenditure/total assets
Assets Tangibility (Tang)	Net property, plant and equipment/total assets
Innovation Investment(RD)	Research and development expenses/total sales
Uniqueness (Unique)	A dummy variable equals one if the firm belongs to industries producing space vehicles, missiles, aircraft, computer, semiconductor, and chemicals, and zero otherwise
Nonproduction cost(SGA)	Selling, general and administrative expenses/total sales
Cash Holdings(Cash)	Cash and short-term investments/total assets
Top Tax Rate(Taxrate)	The top firm tax rate
Depreciation (Depr)	Depreciation and amortization/total assets
Stock Variance (StockVar)	The annual variance of daily stock returns
Bankruptcy Probability(Zscore)	Altman's Z-score (1968)[3]
Debt Rating(Rating)	A dummy variable that equals one if a firm's credit rating is above BBB. The dummy variable is zero if it is lower than BBB.
Stock Returns(StockRet)	Cumulative annual stock returns using monthly raw returns
Market Returns(CrspRet)	Cumulative annual market returns using monthly raw returns
Industry Leverage (Industlev)	The median of firm leverage (TDM)
Industry Growth(industgr)	The median of assets growth (ChgAsset)
Term Spread (Termsprd)	The difference between the 10-year bond returns and the 1-year bond returns
inflation	Expected one-year change in the Consumer Price Index

[3] For emerging countries including Korea, the calculation methodology for Z-score is calculated as follows (Meeampol et al., 2014).

Z = 3.25 + 6.56X1 + 3.26X2 + 6.72X3 + 1.05X4

X1 = (Current Assets – Current Debt) / Total Assets
X2 = Current Profit/Total Assets
X3 = EBIT/Total Assets
X4 = Total Capital/Total Debt

Macroporf	Change in the logarithm of annual firm profits with inventory valuation and capital consumption adjustments
Growth in GDP(Macrogr)	Change in the logarithm of real GDP
Net Payout (Netpay)	(Cash dividends plus purchase of common and preferred stock minus sale of common and preferred stock)/total assets
Size Dummies	A firm is labelled 2, 1, and 0 if the size of the firm (Assets) lies in the upper 30%, middle 40% and lower 30%
Growth Dummies	A firm is labeled high-growth, middle-growth and low-growth if the market-to-book (Mktbk) of firm lies in upper 30%, middle 40% and lower 30%
High-Tech Dummy	A dummy variable that is one if a firm offers technology products and services and zero otherwise

5. Machine learning and leverage

5.1. Predicting performance

Now we conduct and assess the predictive performance of each traditional and machine learning model. We use 2003-2014 as the training period to tune our machine learning model and test out-of-sample performance and MSE over the 2015-2019. In measuring the performance, we estimate the out-of-sample R-squared (R^2_{0S}) and MSE (mean squared error).

In Table 3 and Table 4 below, we present yearly estimates of the out-of-sample R-squared (R^2_{0S}) and MSE respectively. First, it seems evident that the predictive performance of RF excels those of others in most period. The predictability for GBM nearly matches RF's throughout the sample and reaches 0.5272 in 2019. On the other hand, the traditional models, OLS and LASSO are ranging from -0.1052 to 0.40 and -1.2106 to 0.4025 respectively, even showing estimates negative signs in some periods. Thus, our results in overall indicate that machine learning models generally improve on the performance of the linear OLS and LASSO model.

Table 3. Out-of-sample performance (R^2_{0S}) of machine learning versus linear models

	2015	2016	2017	2018	2019	2015~2019
OLS	-0.105230	0.031941	0.401143	0.126228	-1.609892	-0.265298
LASSO	0.006570	0.103022	0.402515	0.191566	-1.210615	-0.127837
RF	0.606779	0.565904	0.523065	0.532053	0.527210	0.552908
GBM	0.591878	0.545335	0.404715	0.537237	0.543981	0.528520

Table 4. Out-of-sample performance (MSE) of machine learning versus linear models

	2015	2016	2017	2018	2019	2015~2019

OLS	0.059471	0.049278	0.031259	0.047088	0.149441	0.068474
LASSO	0.053456	0.045660	0.031187	0.043567	0.126579	0.061035
RF	0.021159	0.022097	0.024895	0.025218	0.027072	0.024195
GBM	0.021961	0.023144	0.031072	0.024939	0.026111	0.025515

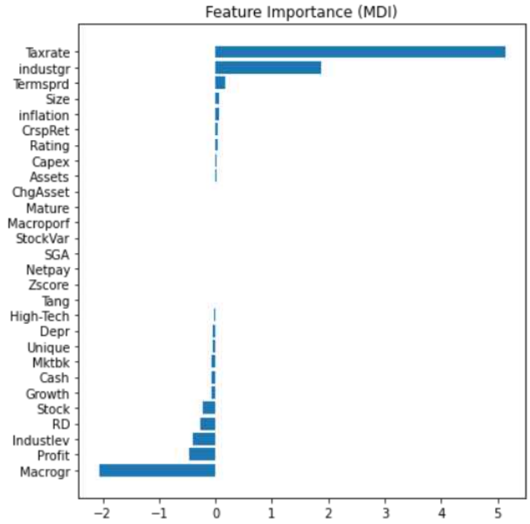
We turn our attention to the estimates of MSE for traditional linear models and machine learning models. As evident in the table, the estimates of RF seem to be the lowest, which ranges from 0.0212 to 0.0271. Again, the estimates of GBM nearly matches those of RF in each year, which ranges from 0.0220 to 0.0311. On the other hand, the estimates of OLS and LASSO range from 0.0313 to 0.1494 and 0.0312 to 0.1266. Interestingly, we find the estimates sharply go up in the last year for each linear model, indicating that their regression line to a set of points is becoming less fitted in later year of our sample.

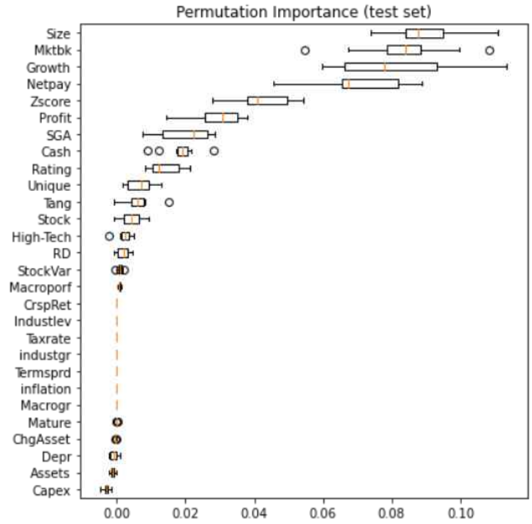
5.2. Variable importance

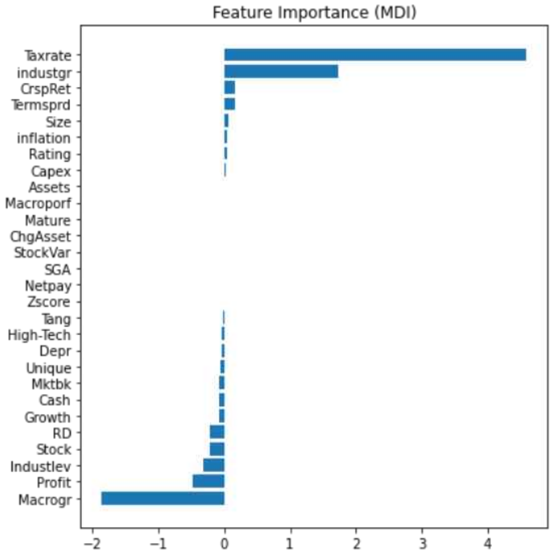
Now we look at the variable importance for the traditional linear models and machine learning models. In Figure 1 below, the Panel A and Panel B represent the variable importance of linear models and LASSO. Subsequently, Panel C and Panel D show the variable importance of RF and GBM respectively. These estimates are obtained from the training sample over 2003-2014 and we use them to calculate the variable importance for 2015. In each panel, the figure to the left shows the variable importance for training set and the figure to the right shows the variable importance for the predicted test set based on the training set data.

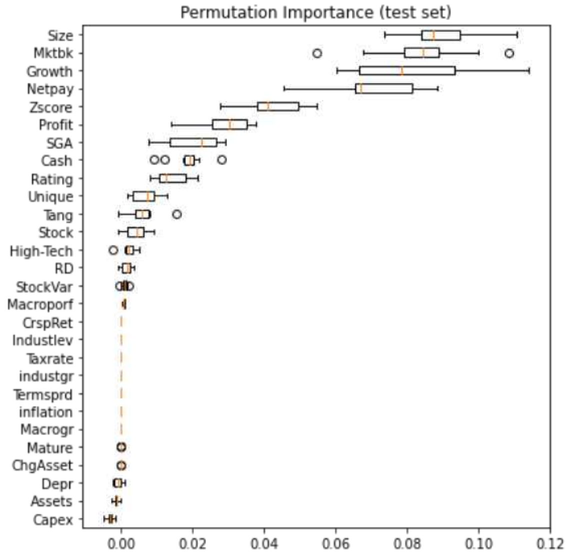
For both linear and LASSO model, the firm-specific factors such as *Size, Mktbk, Growth, Netpay* and *Z-score* are the top five highest importance, indicating that these firm-specific factors have more explanatory power in each model. Rather, the firm-specific variables such as *Capex, Assets, Depr,* and *ChgAsset*, show negative signs, implying their less significant impact in each model.

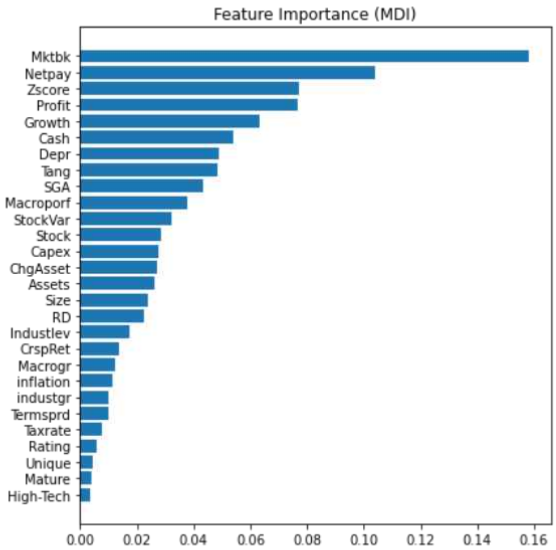
Figure 1. Variable importance from linear and machine learning model (TDM)

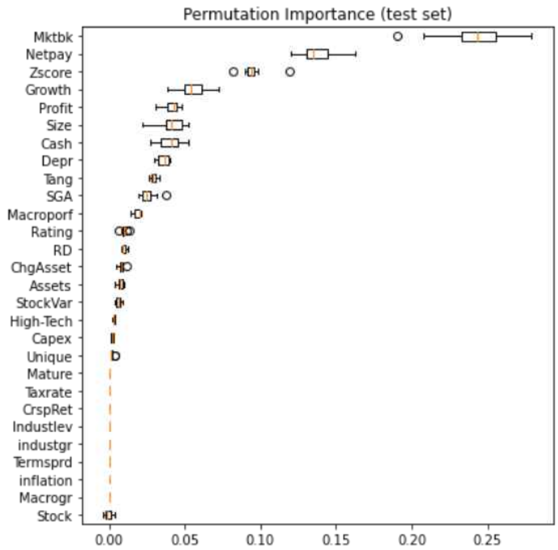

Feature Importance (MDI)
Taxrate
industgr
Termsprd
Size
inflation
CrspRet
Rating
Capex
Assets
ChgAsset
Mature
Macroporf
StockVar
SGA
Netpay
Zscore
Tang
High-Tech
Depr
Unique
Mktbk
Cash
Growth
Stock
RD
Industlev
Profit
Macrogr

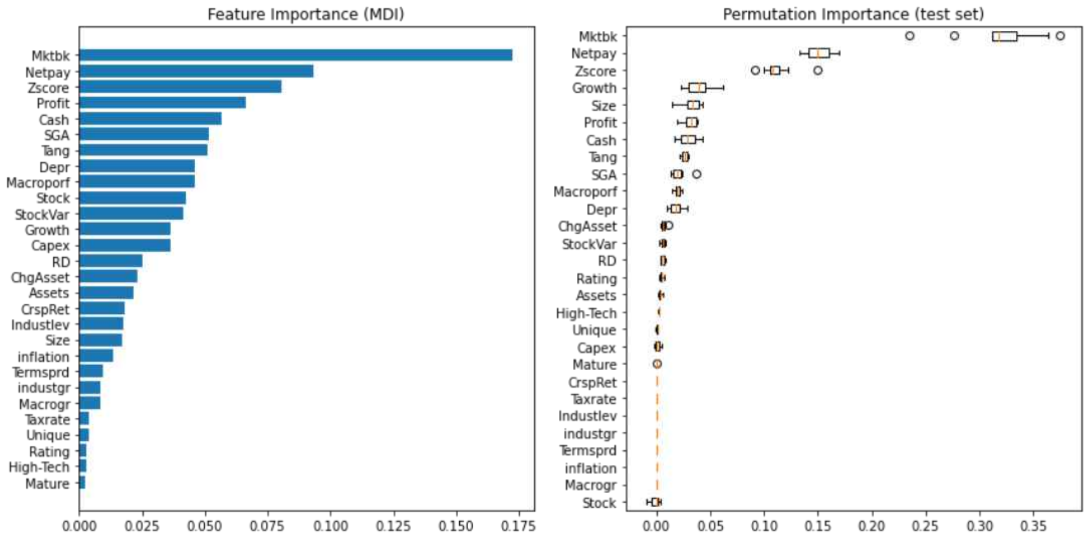

Permutation Importance (test set)
Size
Mktbk
Growth
Netpay
Zscore
Profit
SGA
Cash
Rating
Unique
Tang
Stock
High-Tech
RD
StockVar
Macroporf
CrspRet
Industlev
Taxrate
industgr
Termsprd
inflation
Macrogr
Mature
ChgAsset
Depr
Assets
Capex


Feature Importance (MDI)
Taxrate
industgr
CrspRet
Termsprd
Size
inflation
Rating
Capex
Assets
Macroporf
Mature
ChgAsset
StockVar
SGA
Netpay
Zscore
Tang
High-Tech
Depr
Unique
Mktbk
Cash
Growth
RD
Stock
Industlev
Profit
Macrogr


Permutation Importance (test set)
Size
Mktbk
Growth
Netpay
Zscore
Profit
SGA
Cash
Rating
Unique
Tang
Stock
High-Tech
RD
StockVar
Macroporf
CrspRet
Industlev
Taxrate
industgr
Termsprd
inflation
Macrogr
Mature
ChgAsset
Depr
Assets
Capex


Feature Importance (MDI)
Mktbk
Netpay
Zscore
Profit
Growth
Cash
Depr
Tang
SGA
Macroporf
StockVar
Stock
Capex
ChgAsset
Assets
Size
RD
Industlev
CrspRet
Macrogr
inflation
industgr
Termsprd
Taxrate
Rating
Unique
Mature
High-Tech


Permutation Importance (test set)
Mktbk
Netpay
Zscore
Growth
Profit
Size
Cash
Depr
Tang
SGA
Macroporf
Rating
RD
ChgAsset
Assets
StockVar
High-Tech
Capex
Unique
Mature
Taxrate
CrspRet
Industlev
industgr
Termsprd
inflation
Macrogr
Stock



On the other hand, it seems certain that for both RF and GBM model, the firm-value factors such as *Mktbk, Netpay, Z-score, Profit,* and *Growth* are of the highest variable importance, indicating that their firm-specific factors have more explanatory power in each model. Rather, firm-specific or macroeconomic variables such as *Stock, Macrogr, Inflation, Termsprd* and *Industlev* are of the lowest variable importance, indicating that these factors less more explanatory power in each model. To summarize, our analysis of variable importance for machine learning models demonstrate that the firm-value factors have more explanative power in predictions, while firm-specific or macroeconomic variables have less explanative power.

6. Machine learning and leverage adjustment

6.1. Estimating leverage adjustment speed

Now we conduct further analysis to estimate leverage adjustment speed. In accordance with our theoretical framework (Amini et al., 2021; Fama et al., 2002; Flannery et al., 2006; Kim et al., 2007), firms, in order to rebalance their capital structure, evaluate how quickly to close any gap between actual and target capital structure. We define the gap as $GAP_{i,t} = E(y_{i,t+1}) - y_{i,t}$, and the leverage adjustment model is as follows.

$$\triangle y_{i,t+1} = \lambda GAP_{i,t} + \varepsilon_{i,t+1} \quad \ldots (7)$$

In eq. (7) above, the adjustment speed, λ, allows firms to move partially towards their target leverage during year t. Here, if firm managers have their own target debt ratio and have willingness to reach them, λ should be greater than zero. Put it differently, whenever there exists wedge between target and actual leverage ratio, firms will adjust the ratio accordingly. The Table 5 below presents the adjustment speed estimates for both traditional linear models and machine learning models. We use the dependent variables, TDM and TDA, to observe the GAP, denoted as λ, and half-life. The estimates for linear model, LASSO, RF and GBM are separated into Column 1, Column 2, Column 3 and Column 4 respectively. Further, we present estimates which include firm-fixed effects for each model as presented in Column 5, Column 6, Column 7 and Column 8.

Table 5. The speed of leverage adjustment

	Without firm-fixed effect				With firm-fixed effect			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: TDM	Multi regression	LASSO	RF	GBM	Multi regression	LASSO	RF	GBM
GAP	0.07*** (0.006)	0.08*** (0.006)	0.11** (0.011)	0.11** (0.011)	0.14** (0.013)	0.16** (0.013)	0.47** (0.024)	0.32** (0.022)
Observation	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650
Adjusted R^2_{OS}	0.056	0.057	0.041	0.036	0.056	0.064	0.152	0.091
Half-life in years	3.78	3.69	3.16	3.23	2.81	2.61	1.14	1.64
Panel B: TDA								
GAP	0.05*** (0.008)	0.06*** (0.008)	0.08*** (0.01)	0.08** (0.011)	0.28** (0.02)	0.32** (0.021)	0.41** (0.024)	0.31** (0.023)
Observation	2,650	2,650	2,650	2,650	2,650	2,650	2,650	2,650
Adjusted R^2_{OS}	0.017	0.018	0.022	0.019	0.053	0.067	0.091	0.044
Half-life in years	4.27	4.18	3.62	3.71	1.82	1.63	1.30	1.90

Symbols ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

In Column 1 to Column 4, both machine learning models yield an adjustment speed for 0.11(RF, GBM) for TDM, which barely excels those of linear models, which are at 0.07(Multi-regression) and 0.08(LASSO) respectively. However, when the estimates are translated into half-life, the new estimates in years for machine learning models are 3.16(RF) and 3.23(GBM), which are much shorter than the estimates of traditional linear models which are 3.78(Multi-regression) and 3.69(LASSO). The new estimates for TDA in machine learning models are both shorter than those of traditional linear models. In Column 5 to Column 8, the results with firm-fixed effect are found to be more interesting. The estimates in both machine learning models yield an adjustment speed for 0.47(RF) and 0.32 (GBM) for TDM, which largely excels those of linear models, which is at 0.14(Multi-regression) and 0.16(LASSO) at statistically significant model. The half-life is 1.14 and 1.64 for RF and GBM, which are much shorter than 2.81 and 2.61 for multi-regression and LASSO model. Again, the estimates for adjustment speed for TDA in machine learning models are both greater than those of traditional linear models and accordingly, each half-life is shorter than respective traditional models. All estimates in Table 5 are at statistically significant level.

6.2. Determinants of leverage adjustment speed

We undertake additional analysis in order to capture the importance of explanatory variables for estimating the speed of leverage adjustment. We deliberately select RF model since it shows the most satisfactory performance in both the speed of leverage adjustment and half-life in previous analysis. In particular, we divide the period for training into 2010-2014 and 2010-2018 in Panel A and Panel 2 in Figure 2 below. The variable importance is reported for the last year of the sample, which is 2019.

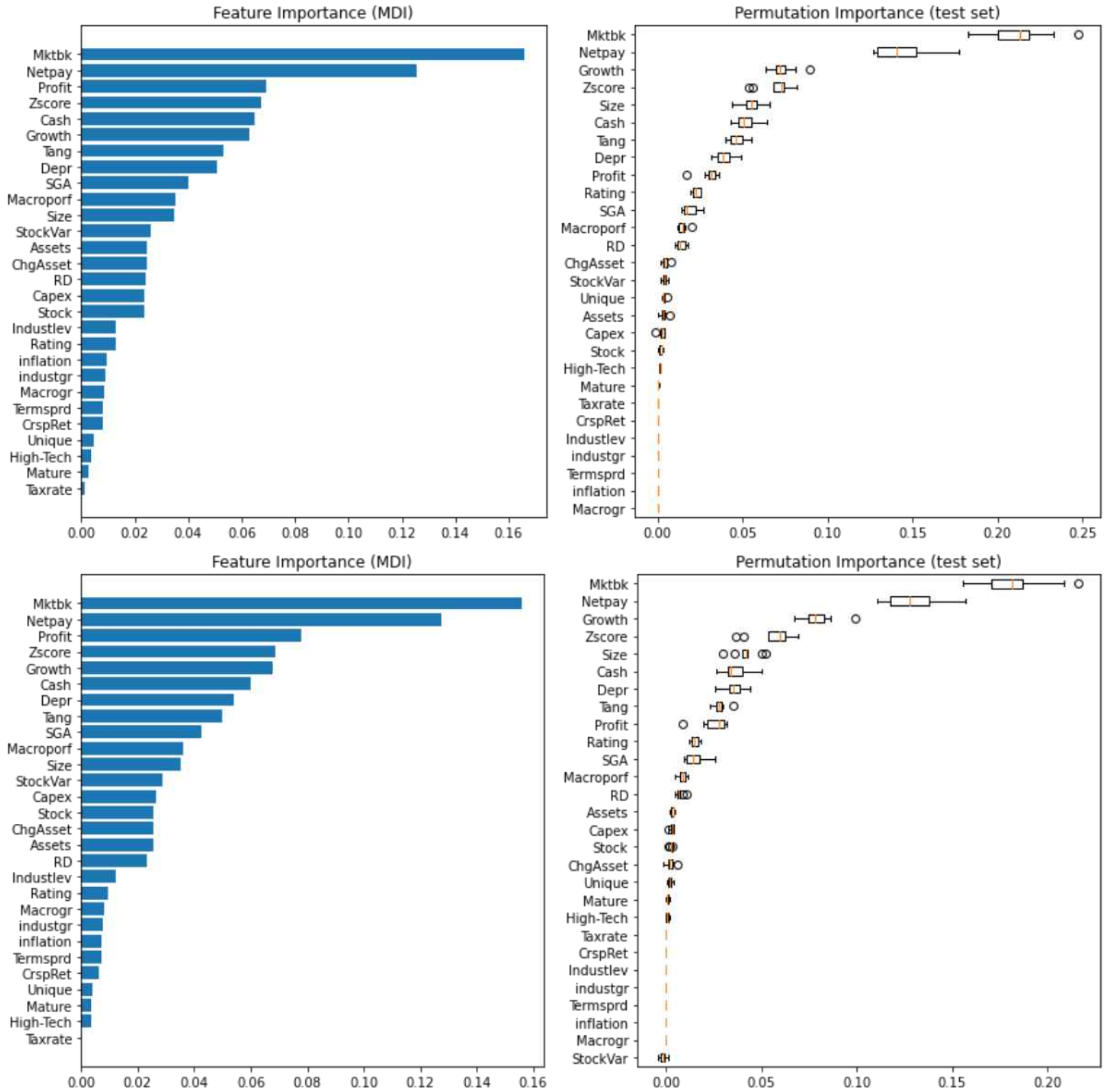


Figure 2. Variable importance from RF for 2010-2014 and 2010-2018 (TDM)

In the Figure above, it is evident that firm-specific factors are mostly of variable importance in both periods. For 2010-2014, *Mktbk, Netpay, Profit, Zscore* and *Cash* are the explanatory variables, while *Mktbk, Netpay, Zscore, Profit* and *Growth* are the explanatory variables for

2010-2018. It is interesting to note that the results of the Figure 2 bears similarity to the results in Figure 1, implying that the firm-value factors are still the most explanatory variables in machine learning models.

7. Additional analysis

7.1. Leverage adjustment speed for chaebols and non-chaebols

Now, we conduct additional analysis in order to find out the heterogeneity effect for firms affiliated to *chaebols*. In last decades, there has been a wide debate and discussion over the determinants of firm debt in Korea (Lee et al, 2000; Kim et al, 2007; Yoon et al., 2016). More than often, profitability, firm size and growth rate are well-known as common determinants in financing decision for Korean firms. In particular, Lee et al (2000) notes that *chaebols* have much higher leverage than non-*chaebols* firms in Korea, implying their less flexibility to adjust to new leverage ratio and making financing decision when in recession and credit crunch. Here, it seems reasonable to argue that in comparison with non-*chaebols*, it would take more speed of leverage adjustment for *chaebols*, leading to the increase of half-life in years. In fact, we find striking difference across *chaebols* and non-*chaebols*, as presented in the Table 6 and Table 7 below, respectively.

Table 6. The speed of leverage adjustment (chaebols)

	Without firm-fixed effect				With firm-fixed effect			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: TDM	Multi regression	LASSO	RF	GBM	Multi regression	LASSO	RF	GBM
GAP	0.09*** (0.000)	0.10*** (0.000)	0.08*** (0.000)	0.08 *** (0.000)	0.19*** (0.000)	0.22*** (0.000)	0.31*** (0.000)	0.22*** (0.000)
Observation	562	562	562	562	562	562	562	562
Adjusted R^2_{0S}	0.106	0.103	0.029	0.024	0.259	0.269	0.250	0.214
Half-life in years	7.32	6.93	8.19	8.23	3.33	2.79	1.88	2.86
Panel B: TDA								
GAP	0.04*** (0.004)	0.04*** (0.003)	0.03* (0.066)	0.03 (0.137)	0.25*** (0.000)	0.27*** (0.000)	0.25*** (0.000)	0.14*** (0.001)

Observation	562	562	562	562	562	562	562	562
Adjusted R^2_{OS}	0.013	0.013	0.005	0.002	0.178	0.183	0.154	0.124
Half-life in years	18.16	17.92	21.32	25.58	2.40	2.24	2.40	4.57

Symbols ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 7. The speed of leverage adjustment (non-chaebols)

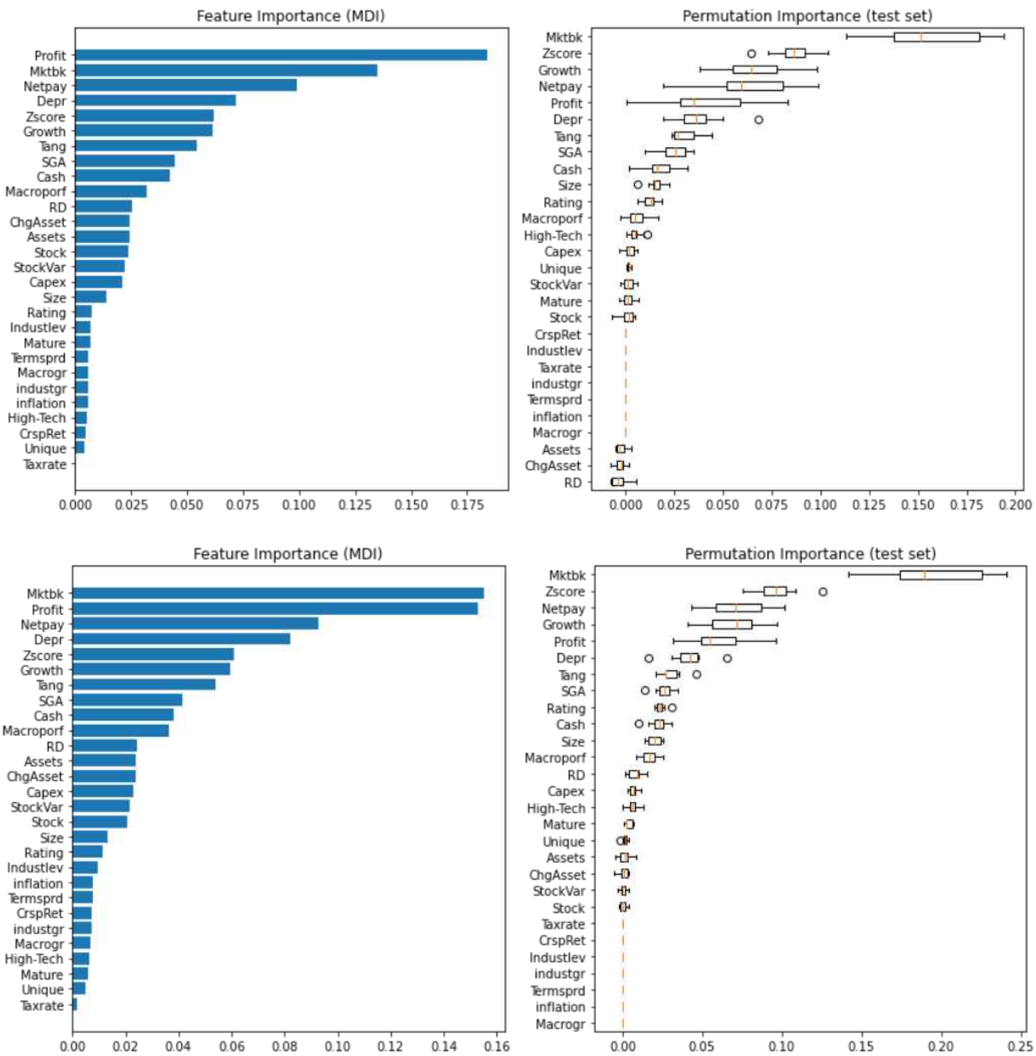
	Without firm-fixed effect				With firm-fixed effect			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: TDM	Multi regression	LASSO	RF	GBM	Multi regression	LASSO	RF	GBM
GAP	0.07*** (0.000)	0.08*** (0.000)	0.12*** (0.000)	0.12*** (0.000)	0.14*** (0.000)	0.16*** (0.000)	0.48*** (0.000)	0.36*** (0.000)
Observation	2088	2088	2088	2088	2088	2088	2088	2088
Adjusted R^2_{OS}	0.050	0.052	0.045	0.040	0.022	0.030	0.135	0.082
Half-life in years	9.50	8.80	5.42	5.68	4.75	4.09	1.05	1.54
Panel B: TDA								
GAP	0.05*** (0.000)	0.06*** (0.000)	0.09*** (0.000)	0.08*** (0.000)	0.26*** (0.000)	0.31*** (0.000)	0.43*** (0.00)	0.31*** (0.000)
Observation	2088	2088	2088	2088	2088	2088	2088	2088
Adjusted R^2_{OS}	0.017	0.018	0.026	0.020	0.035	0.050	0.089	0.036
Half-life in years	12.85	11.69	7.28	8.13	2.30	1.90	1.25	1.86

Symbols ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

In Table 6, the results for both linear models and machine learning models are contrasting to the analysis in Table 5. First, the leverage adjustment speed of machine learning models is slower than those of traditional linear models without firm-fixed effect. However, even if firm-fixed effect is included, the estimates of machine learning models barely excel those of traditional linear models. In contrast, however in the Table 7 for non-*chaebols*, the results for both linear models and machine learning models are similar to the analysis in Table 5 in that (1) the leverage adjustment speed of machine learning models is faster than those of traditional linear models with, or without firm-fixed effects, and (2) the half-life of machine learning models is shorter than those of traditional linear models. Interestingly, the leverage adjustment speed of non-*chaebols* when measured in RF with firm-fixed effect is 0.48 and 0.43 which is much faster than those of *chaebols* which are at 0.31 and 0.25, which lead to large difference in half-life in years. Thus, it is reasonable to conclude that our analysis well conforms to the characteristics of *chaebols* defined by previous researches (Lee et al, 2000; Yoon et al, 2005).

7.2. Variable importance for chaebols and non-chaebols

Similar to previous section, we conduct random forest model to analyze and rank the impact of factors on the speed of leverage adjustment. The figure 3 below plots the importance of explanatory variables for estimating the speed of adjustment using random forest (RF) model for *chaebols* (Panel A and Panel B) and non-*chaebols* (Panel C and Panel D). The estimates are obtained from training the model over the period 2010-2015 and 2010-2018. Accordingly, the variable importance is reported for the last year of the sample, 2019.



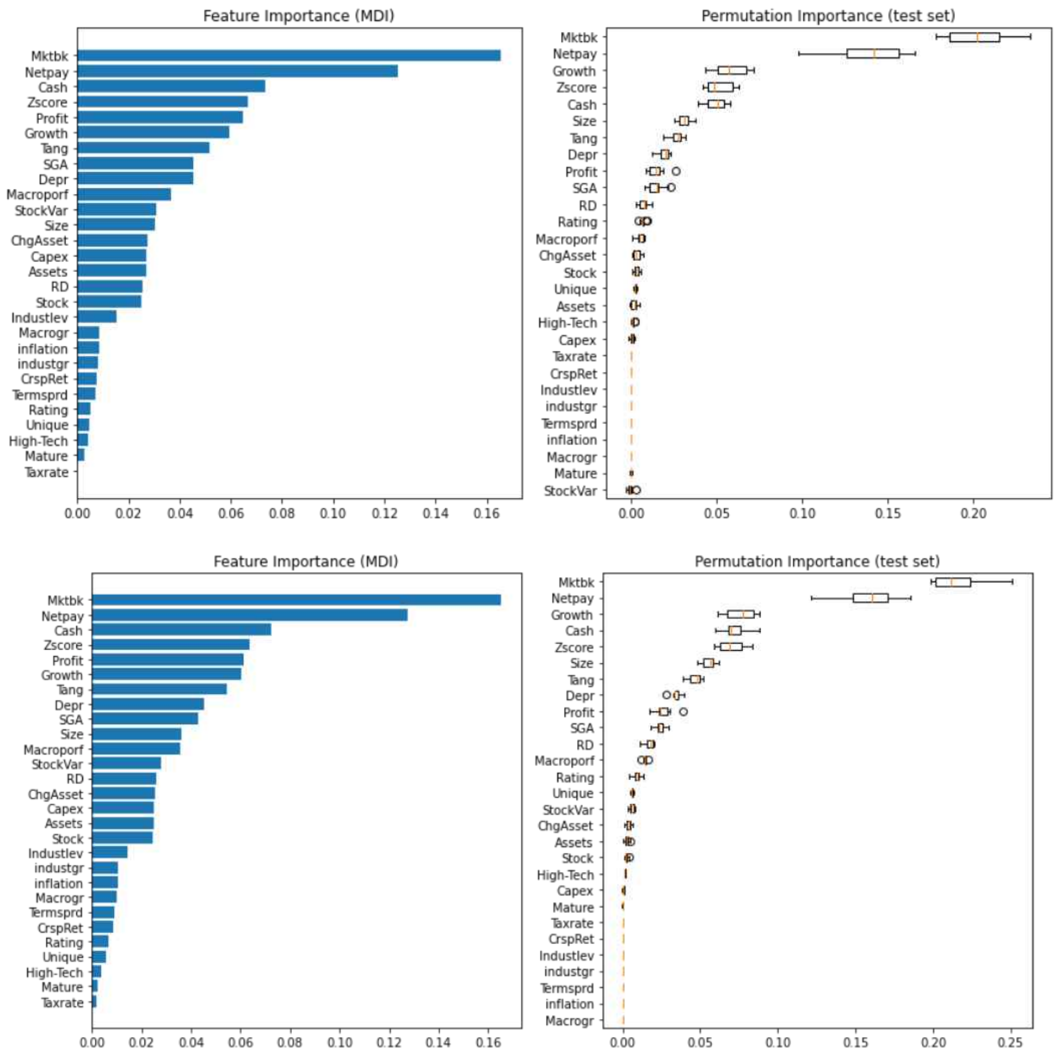


Figure 3. Variable importance from RF for 2010-2015 and 2010-2018 (TDM)

For the *chaebols* in Panel A and Panel B in the Figure 3, it appears that *Profit* seems to be one of the most determinant in financing decision for the affiliates. This is in accordance with the previous researches in that profitability of *chaebols* is a major concern for investors and capital markets. In fact, the firm value of *chaebols* has been assessed on the basis of profitability, thanks to the long support of government. Since the 1997 Asian financial crisis, *chaebols* have become a very profitable firms with less over-investment despite fewer tax perks (Baek et al, 2004; Lee et al, 2008). In contrast, for non-*chaebols*, the firm-specific factors such as *Mktbk, Netpay, Cash, Z-score* and *Growth* are found to be the most determinants in their financing decisions, which is in accordance with the results in the analysis for all firms in the Figure 1. This finding bears significant importance as follows. Not only does machine learning model captures the variable importance more effectively than other traditional models, but also, it better explains the characteristics of the financing decision of *chaebols*, which is unique to Korean business environment.

7.3. Leverage adjustment speed for chaebols and non-chaebols

Lastly, we present leverage adjustment speed and half-year for every model with the different dependent variables, TDM, TDA, LDM, and LDA, for all firms, *chaebols* and non-*chaebols* In Table 6 below. In general, the leverage adjustment speed of non-*chaebols* is found to be faster than that of *chaebols* in most of the leverage variables. Exceptionally, however, for LDM with firm-fixed effect, the leverage adjustment speed of *chaebols* is found to be faster than that of non-*chaebols*, which translate into shorter half-year. We speculate that this well reflects the financial behavior of *chaebol*-affiliated firms: A partial adjustments to their long-run targets. The potential explanation is that the interaction between the financial decisions and long-run financial targets – their tendency to cash holding, for example - will allow variations in the speed of adjustments over time (Jalilvand and Harris, 1984).

Table 6. Leverage adjustment speed and half-year for all firm, chaebols and non-chaebols

			All firm		Chaebol		Non-chaebol	
	Model	Firm-fixed effect	Leverage adjustment speed	half-year	Leverage adjustment speed	half-year	Leverage adjustment speed	half-year
TDM	linear	No	0.0736	9.0694	0.0903	7.3230	0.0703	9.5038
	lasso		0.0788	8.4536	0.0952	6.9274	0.0758	8.7952
	RFR		0.1130	5.7806	0.0812	8.8188	0.1201	5.4184
	GBM		0.1057	6.2040	0.0807	8.2328	0.1150	5.6756
	linear	Yes	0.1425	4.5098	0.1880	3.3276	0.1358	4.7508
	lasso		0.1639	3.8715	0.2199	2.7913	0.1559	4.0889
	RFR		0.4566	1.1363	0.3087	1.8777	0.4827	1.0515
	GBM		0.3156	1.8278	0.2152	2.8607	0.3628	1.5383
TDA	linear	No	0.0528	12.7711	0.0375	18.1594	0.0525	12.8498
	lasso		0.0563	11.9686	0.0379	17.9190	0.0576	11.6929
	RFR		0.0836	7.9385	0.0320	21.3196	0.0908	7.2788
	GBM		0.0774	8.6078	0.0267	25.5843	0.0817	8.1327
	linear	Yes	0.2818	2.0943	0.2512	2.3964	0.2602	2.2996
	lasso		0.3213	1.7883	0.2663	2.2386	0.3060	1.8972
	RFR		0.4144	1.2952	0.2510	2.3978	0.4263	1.2474
	GBM		0.3059	1.8981	0.1408	4.5676	0.3113	1.8589
LDM	linear	No	0.2046	3.0278	0.1865	3.3573	0.2097	2.9446
	lasso		0.2081	2.9703	0.1899	3.2918	0.2514	2.8567
	RFR		0.2503	2.4058	0.2380	2.5497	0.2414	2.5088
	GBM		0.2539	2.3667	0.2596	2.3066	0.2560	2.3444
	linear	Yes	0.7814	0.4559	0.8771	0.3306	0.7393	0.5156
	lasso		0.8283	0.3934	0.8907	0.3131	0.8144	0.4115
	RFR		0.8333	0.3869	0.8549	0.3591	0.8543	0.3598
	GBM		0.8167	0.4086	0.8397	0.3786	0.8392	0.3793
LDA	linear	No	0.2463	2.2452	0.1663	3.8118	0.2673	2.2287
	lasso		0.2500	2.4092	0.1811	3.4696	0.2680	2.2217
	RFR		0.3008	1.9373	0.2440	2.4781	0.3199	1.7979
	GBM		0.3010	1.9358	0.2382	2.5471	0.3219	1.7842
	linear	Yes	0.8539	0.3604	0.7240	0.5384	0.8700	0.3397
	lasso		0.8691	0.3409	0.8290	0.3925	0.8735	0.3352
	RFR		0.8707	0.3388	0.8671	0.3435	0.8905	0.3134
	GBM		0.8420	0.3757	0.7863	0.4491	0.8632	0.3485

8. Conclusion

The results of our analysis are as follows. First, our machine learning approach achieves substantial gains in estimating target leverage. For example, R^2_{0S} statistics for RF and GBM are found to be greater than those for OLS and LASSO over the testing period (2015-2019), with or without firm-fixed effect. Further, it seems clear that RF and GBM outperform linear models, leading to lower error every year in the testing period, as measured in with or without firm-fixed model. Second, our results demonstrate that the leverage adjustment speed of machine learning models is much faster than those of traditional linear models, which translate into much lower half-life in years. Third, we find heterogeneity effect for *chaebols* and non-*chaebols* in leverage adjust speed. For instance, the leverage adjustment speed for non-*chaebols* is much faster than that of *chaebols* in most model used, including machine learning and traditional models with or without firm-fixed effect. This could be due to debt-dependent characteristic of *chaebols* which are known to have extremely high leverages when compared to non-c*haebols* and firms in other countries (Lee, 2000).

The contribution of our research is as follows. First, to our knowledge, this study is first research to employ machine learning methodologies in analyzing target leverage for Korean firms. Second, we verified how machine learning models, RF and GBM, better estimate leverage adjustment speed and half-life year than traditional models, thereby better conforming to the trade-off theory. Lastly, we find the effect of heterogeneity between *chaebols* and non-*chaebols*, on *both* leverage adjustment speed and half-life year, proving that the characteristics of *chaebols*, high leverages, are still identified and considered in machine learning models.

References

Altman, N. and Krzywinski, M., 2017, Ensemble methods: bagging and random forests, *Nat. Methods* 14, pp. 933–934.

Baek, J.S., Kang, J.K., Park, K.S., 2004, Corporate governance and firm value: evidence from the Korean financial crisis, *Journal of Financial Economics* 71, pp.265-313.

Bong-han Yoon, Jae-Yong Oh, 2005, Korean Case Studies on Corporate Governance & Firm`s Performance, Value and Market Returns, *Asian Pacific Journal of Financial Studies*, Vol34, pp.227-263

Choi Yun Yee, 2015, A Study on the influences of Debt Ratio before and after the Financial Crisis on Firm's value, *Journal of Industrial Economics and Trade*, Vol.39, no.1, pp.35-59

Cristina Aybar-Aria, Alejandro Casino-Martínez and José López-Gracia, 2012, "On the adjustment speed of SMEs to their optimal capital structure," *Small Business Economics* 39, pp.977–996.

Efron, B. and Tibshirani, R.J., 1994,. An Introduction to the Bootstrap, *Chapman and Hall/CRC, UK*.

Erwan Morellec, 2004, "Can Managerial Discretion Explain Observed Leverage Ratios?" *The Review of Financial Studies 17*, pp.257-294.

Eugene F. Fama and Kenneth R. French, 2002, Testing Trade-Off and Pecking Order Predictions about

Dividends and Debt, *The Review of Financial Studies* 15, pp.1-33.

Eugene F. Fama and James D. MacBeth, 1973, Risk, return, and equilibrium: empirical tests, *The Journal of Political Economy* 81, pp.607-636.

Heejung Yeo , 2018, Target Leverage and Determinants of Leverage in Shipping Companies, *Korean trade review* vol.43, no.2 , pp.181 - 204

Jinsu Kim, 2010, Financing Market Accessibility and Capital Structure Adjustment Speed, *Asian Review of Financial Research*, pp.89-120.

Minshik Shin, Kim Soo Eun, 2011, "The Relationship between Ownership Structure and Market Value of Firms," *Institute for Insurance and knowledge Myong Ji University* vol.9, no.2, pp. 129-157

Murray Z. Frank and Vidhan K. Goyal Frank, March 2004, The effect of market conditions on capital structure adjustment, *Finance Research Letters* 1, pp.47-55.

Giovanni S.F. Bruno, "Approximating the bias of the LSDV estimator for dynamic unbalanced panel data models," *Economics letters* 87, June 2005, 361-366.

Michael C. Jensen, Agency Costs of Free Cash Flow, May 1986, Corporate Finance, and Takeovers", *The American Economic Review* 76, pp.323-329.

Michael C. Jensen and William H. Meckling, October 1976, Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics* 3, pp.305-360.

Jalilvand, A. and R. Harris, 1984, Corporate behavior in adjusting to capital structure and dividend targets: An econometric study, *Journal of Finance* 39, pp.127-145.

Jerome H. Friedman, 29, October 2001, Greedy Function Approximation: A Gradient Boosting Machine, *The Annals of Statistics,* pp.1189-1232.

John R. Graham and Campbell R. Harvey, May 2001, The theory and practice of corporate finance: evidence from the field, *Journal of Financial Economics* 60, pp.187-243.

Julio Pindado and Alberto de Miguel, December 2001, Determinants of Capital Structure: New Evidence from Spanish Panel Data, *Journal of Corporate Finance* 7, pp.77-99.

Lakshmi Shyam-Sunder and Stewart C. Myers, February 1999, Testing static tradeoff against pecking order models of capital structure, *Journal of Financial Economics* 51, pp.219-244.

Lee, J.W. , Lee , Y.S., Lee, B.S., May 2003, The Determination of Corporate Debt in Korea, *Asian Economic Journal,* pp.333-356

Lee, K., Kim, J.Y., Lee, O., 2010. Long-term evolution of the firm value and behavior of business groups : Korean chaebols between weak premium, strong discount, and strong premium, *Journal of The Japanese and International Economies* 24 (3), pp.412-440.

Leo Breiman, August 1996, Bagging predictors, *Machine Learning* 24, pp.123–140.

Leo Breiman, "Random Forests," *Machine Learning* 45, October 2001, pp.5–32.

Malcolm Baker and Jeffrey Wurgler, "Market Timing and Capital Structure," *The Journal of Finance* 57, December 2002, pp.1-32.

Mark Flannery and Kasturi P. Rangan, “Partial adjustment toward target capital structures,” *Journal of Financial Economics* 79, March 2006, pp.469-506.

Murray Z. Frank and Vidhan K. Goyal, Spring 2009, “Capital Structure Decisions: Which Factors Are Reliably Important?” *Financial Management* 38, pp.1-37

Paul D.Childs, David C.Mauer and Steven H.Ott., June 2005, Interactions of corporate financing and investment decisions: The effects of agency conflicts, *Journal of Financial Economics* 76, pp.667-690.

Pando Son· Seung-Tae Son, 2008, Capital Structure's Mean-Reversion and Long-Term Equilibrium, *Korea Financial Management Association,* vol.25, no.3, pp. 33-78

RenéM.Stulz, 1990, Managerial Discretion and Optimal Financing Policies, *Journal of financial Economics* 26, pp.3-27.

Robert Tibshirani, July 1990, Regression Shrinkage and Selection via the Lasso, *Journal of the Royal Statistical Society* 58, pp.267-288.

Sasivimol Meeampol, Polwat Lerskullawat, Ausa Wongsorntham, Phanthipa Srinammuang, Vimol Rodpetch and Rungsimaporn Noonoi, “Applying Emerging Market Z-Score Model To predict Bankruptcy: A case study of Listed companies In The Stock Exchange Of Thailand (Set),” Management, Knowledge and learning International conference 2014.

Shahram Amini, Ryan Elmore, Ozde Oztekim and Jack Strauss, October 2021, Can machines learn capital structure dynamics?” *Journal of Corporate Finance* 70, pp.1–22.

Seung-Tae Son, Yoon-Goo Lee, 2007, Determinants of Capital Structure in KOSDAQ Firms, *The Korean journal of financial management* vol.24 no.1 , pp.109 – 147.

Stewart C. Myers, July 1984, The Capital Structure Puzzle, *The Journal of Finance* 39, pp.574-592.

Strobl, C., Boulesteix, A., Kneib, T., Augustin, T. and Zeileis, A., “Conditional variable importance for random forests,” *BMC Bioinform* 9, July 2008, 307–317.

Yeongrae Kim, Pilgyu Jongbeom Choi, 2007, Determinants of capital structure and partial adjustment toward target capital structures, *The Korean Finance Association*』

YeonHee Park, Yosoon Choi, Hansu Kim, “The association between discretionary accruals and bond yields,” 『 Korean Accounting Association 』2011, vol.20, no.4, pp. 35-56.

Yoon Bohyun, Jeong Hwan Lee, Ha, Joon , Sam Ho Son , 2016, The Speed of Adjustment toward Target Leverage in Korean Firms: Based on New Models, *Regional industry review* vol.39, no.2, pp. 55-84.