Journal of Corporate Finance 70 (2021) 102073

Contents lists available at ScienceDirect

Journal of Corporate Finance

journal homepage: www.elsevier.com/locate/jcorpfin

Can machines learn capital structure dynamics?☆

Shahram Amini[a,*], Ryan Elmore[b], Özde Öztekin[c], Jack Strauss[a]

[a] Daniels College of Business, Finance Department, University of Denver, United States
[b] Daniels College of Business, Business Information and Analysis Department, University of Denver, United States
[c] College of Business, Finance Department, Florida International University, United States

ARTICLE INFO

JEL classification:
G17
G30
G32
C10
C50

Keywords:
Machine learning
Target leverage
Speed of leverage adjustment
Financing actions

ABSTRACT

Yes, they can! Machine learning models predict leverage better than linear models and identify a broader set of leverage determinants. They boost the out-of-sample R^2 from 36% to 56% over OLS and LASSO. The best performing model (random forests) selects market-to-book, industry median leverage, cash and equivalents, *Z*-Score, profitability, stock returns, and firm size as reliable predictors of market leverage. More precise target estimation yields a 10%–33% faster speed of adjustment and improves prediction of financing actions relative to linear models. Machine learning identifies uncertainty, cash flow, and macroeconomic considerations among primary drivers of leverage adjustments.

1. Introduction

How firms choose their capital structure and whether they actively move towards a target leverage ratio are among the central issues in corporate finance. Despite extensive debate and research over decades (Myers, 1984; Titman and Wessels, 1988; Harris and Raviv, 1991; Rajan and Zingales, 1995, among others), the factors that drive the cross-sectional and time-series variations in capital structure remain a puzzle (Lemmon et al., 2008; Rauh and Sufi, 2012; DeAngelo and Roll, 2015). How to reliably estimate leverage targets and the speed of adjustment to those targets are also largely unresolved. Why does previous research produce such disparate estimates of firm leverage adjustment speed? What variables robustly determine the desired level of leverage and the convergence rate to target leverage? Can leverage determinants and estimated targets consistently predict firms' financing actions?

Our paper employs machine learning models including random forests (RF), gradient boosting machines (GBM), neural networks (NNET), generalized additive models (GAM), and least absolute shrinkage and selection operator (LASSO) models to answer these questions. We exploit a large sample comprising 128,417 firm-year observations from 1972 to 2018 and use a training and cross-validation period from 1972 to 2000 to estimate parameters and tune these models. We then assess out-of-sample R-squared (here-

☆ We are indebted to Don Bowen, Conrad Ciccotello, Andrew Detzel, Mark Flannery, Todd Gormley, Chris Hughen, Lingfei Kong, Sattar Mansi, Christopher Parmeter, Jay Ritter, Johan Sulaeman, Sheridan Titman, Tracy Xu, Emmanuel Yimfor, Guofu Zhou, participants at the 37th International Conference of the French Finance Association, the 2020 Financial Management Association Conference, the 2020 Boca Corporate Finance and Governance Conference, the 40th International Symposium on Forecasting, and seminar participants at Virginia Tech and University of Denver for their helpful comments and suggestions.

* Corresponding author.
E-mail addresses: shahram.amini@du.edu (S. Amini), ryan.elmore@du.edu (R. Elmore), ooztekin@fiu.edu (Ö. Öztekin), jack.strauss@du.edu (J. Strauss).

https://doi.org/10.1016/j.jcorpfin.2021.102073
Received 29 April 2021; Received in revised form 10 August 2021; Accepted 14 August 2021
Available online 18 August 2021
0929-1199/

after R^2_{OS}) and out-of-sample root mean squared forecast error (RMSE) from 2001 to 2018 to evaluate the predictive power of our machine learning models relative to conventional linear models in measuring target leverage and analyzing its determinants. Next, we estimate the speed of adjustment to target leverage and examine whether machine learning models can correctly predict observed firms' financing actions. We particularly emphasize RF results because RF's ensemble nature and random element correct for decision trees' habit of overfitting to their training set even when the process is nonlinear or involves complex, high-order interaction effects (Strobl et al., 2008; Altman and Krzywinski, 2017).

Our machine learning approach achieves substantial gains in estimating target leverage. R^2_{OS} statistics are 39.8% and 36.2% for OLS and LASSO over the testing period (2001–2018), compared to R^2_{OS} statistics of 56.2%, 55.3%, 52.0%, and 46.6% for RF, GBM, NNET, and GAM, respectively. On an annual basis, RF's R^2_{OS} ranges from 45%–58%, while R^2_{OS} for OLS and LASSO ranges from 28%–44% and 30%–39%, respectively. RF and GBM particularly outperform linear models and lead to lower errors every year in the testing period. Moreover, machine learning models significantly reduce RMSE. For example, RF leads to a 14.7% decline in RMSE, compared to OLS. The differences in RMSE are significant at the 1% level.

We then turn to identifying the most important determinants of corporate leverage. Prior studies examining associations between leverage and various factors (e.g., Frank and Goyal, 2009; Öztekin, 2015) implement *in-sample* tests to assess the importance of potential leverage determinants. In contrast, our paper implements *out-of-sample* tests to exploit complex and high-dimensional patterns in leverage behavior and quantify the relative importance of its determinants. Our primary analyses confirm that machine learning models draw information from a broader set of characteristics than existing studies. RF and GBM both show that market-to-book, industry median leverage, cash, Altman's (1968) *Z*-Score, profitability, stock returns, and firm size are the primary determinants of market leverage. LASSO, which exhibits poor out-of-sample performance, selects industry median leverage and cash as determinants of corporate leverage. Overall, while our results provide evidence for the prominence of core leverage factors identified in related studies (Frank and Goyal, 2009; Öztekin, 2015), we demonstrate the importance of additional factors, including bankruptcy risk (*Z*-Score), stock market conditions (cumulative annual stock returns and annual stock variance), liquidity (cash), and discretionary expenditures (R&D, and selling, general, and administrative costs). Results for book leverage are similar, except cash and Z-Score rank more important for determining book leverage. We find no evidence of changes in firm determinants or a breakdown in predictability over time. The same variables selected by RF are consistently important in predicting leverage. These findings challenge the conventional wisdom that the standard set of firm and macroeconomic determinants has limited ability to explain firms' leverage choices (e.g., Lemmon et al., 2008; Frank and Goyal, 2009).

Next, we ask whether supplementing a standard linear leverage model with prevalent nonlinear specifications and a full set of interactions between the included covariates can achieve similar improvements in predictability as machine learning models. Our tests show that machine learning models detect nonlinearities beyond the typical nonlinear specifications. Adding common nonlinear forms, such as squared, cubed, or lagged terms or their various combinations, to an OLS or LASSO model is not sufficient to achieve the improvements obtained from machine learning. RF continues to outperform these augmented models and generates substantially higher R^2_{OS}.

The hypothesis that firms pursue leverage targets has received mixed support. We add to this debate by using superior target estimates from machine learning and documenting that firms converge to their target leverage faster than is generally understood in the literature. We find that the adjustment speed rises from 16% for LASSO to 22% for RF using a specification that includes firm, industry, and macroeconomic features. Our estimates show that the typical firm closes up to one-third of the leverage gap annually using book leverage, and one-half of the gap using market leverage. These results indicate that estimates of leverage adjustment speed from linear models likely exhibit downward bias due to a failure to capture nonlinearities and interactions between leverage and its determinants.

Numerous papers investigate firm- and country-level determinants of the adjustment speed (e.g., Faulkender et al., 2012; Öztekin and Flannery, 2012; Çolak et al., 2018). However, these studies typically focus on a single determinant in isolation. Hence, it is unclear which determinants matter more for capital structure rebalancing behavior. We are the first to use machine learning to evaluate the relative ranking of a large set of predictors for leverage rebalancing and document novel results. We find that the most prominent drivers of adjustment speed are uncertainty, cash flow, and macroeconomic considerations. On the other hand, the uniqueness and the nature of a firm's assets or a firm's industry are the least important factors for determining adjustment speed.

Theory posits and empirical work partially supports the notion that firms not at their leverage targets change their debt and/or equity to adjust their leverage ratios (e.g., Faulkender et al., 2012; Frank and Shen, 2019). We examine whether target leverage estimates obtained from machine learning models have implications for corporate decisions in the debt and equity markets. We show that, compared to conventional models, the financing actions predicted by machine learning models are more consistent with the predictions of theory. Leverage targets produced by standard linear models fail to correctly predict the observed financing patterns for over or underleveraged firms. The predictions are generally inconsistent with those of theories of capital structure, undermining their empirical credibility. Machine learning models, on the other hand, predict financing actions including equity and debt issuance, stock buybacks, and debt retirement with remarkable accuracy and high statistical significance. This, in turn, implies that firms actively aim towards the target to close their leverage deviation, congruent with theory.

Finally, we investigate whether the superior predictive performance of machine learning methods extends into subsamples constructed based on different firm attributes, including size, growth opportunities, information opacity, and major capital restructuring events. Machine learning models continue to predict leverage more accurately for firms with different characteristics.

Why apply machine learning to examine capital structure? First, the measurement of target leverage and speed of adjustment is fundamentally a problem of prediction. However, leverage adjustments are discrete, infrequent, and large, posing significant challenges in estimating the targets. Consequently, machine learning models that fit complex, nonlinear functional forms can be better suited to estimating capital structure dynamics as these models generally provide significant improvements in out-of-sample forecasting. This is particularly appealing because accurately predicting unobservable leverage targets is essential for estimating adjustment speed, predicting firms' financing patterns (Frank and Shen, 2019), and assessing the empirical validity of capital structure theories (Frank and Goyal, 2009).

Second, nonlinearities and interactions are relevant as well as intrinsic to the relation between leverage and its determinants, and machine learning algorithms search for such patterns automatically. Graham and Leary (2011) find that common variables used to determine capital structure have nonlinear relations with dependent variables, and yet few empirical tests explicitly model these nonlinearities when predicting target leverage or quantifying a determinant's importance. To highlight the relevance of accounting for nonlinear relations between possible determinants and leverage, Fig. 1 fits (marginal) smooth functions using nonparametric splines to commonly used covariates relative to leverage (a generalized additive model approach). The plots in this figure illustrate that the relationship between these determinants and leverage manifests in a nonlinear manner.[1] As a result, machine learning models should be better able to predict target leverage since these models are equipped to handle nonlinear patterns and complex interactions between predictors that are not identified a priori. Such conditions are endemic in a firm's financing decisions (Childs et al., 2005).

Third, in the presence of a large number of covariates (Harris and Raviv, 1991), prior studies use standard model selection techniques to parse through the pool of variables to identify the most salient determinants of leverage (e.g., Frank and Goyal, 2009). However, using conventional models may lead to misidentification of relevant leverage determinants if nonlinearities and complex interactions are ignored. In contrast to these methods, which are likely to suffer from overfitting and redundant variation among explanatory variables, machine learning techniques can more accurately classify an extensive number of determinants. Machine learning estimators specialize in dimension reduction techniques as they are equipped with efficient algorithms capable of searching among the entire model space and removing redundant variation among potentially highly correlated predictors (Gu et al., 2020).

Since our study confirms the success of machine learning algorithms in predicting target leverage and financing actions in a highly nonlinear, discrete environment, we anticipate that these algorithms will be beneficial in areas of finance where various frictions such as high transaction costs and irreversibilities prevent firms from continuously and dynamically adjusting their actions. These frictions tend to create complex, nonlinear patterns between firm decisions or outcomes and their determinants. Consequently, machine learning models that are, by design, better equipped to forecast intermittent patterns in data are more appropriate choices under these circumstances. Therefore, machine learning techniques have great potential for improving studies examining investments in physical capital and R&D, payouts, and cash holdings as these policies generally involve lumpy, discontinuous adjustments.

The remainder of this paper is structured as follows. Section 2 highlights our contributions. Section 3 outlines the empirical methodology. Section 4 reports the sample and summary statistics. Section 5 presents prediction results for leverage targets. Section 6 examines firms' adjustment speed and financing actions. Section 7 offers robustness tests, while Section 8 concludes.

2. Contributions to the literature

The trade-off theory maintains that firms have a value-maximizing (optimal) leverage, which they determine by assessing the trade-off between the benefits and costs of raising debt (Myers, 1977; Jensen, 1986). The pecking-order theory does not assume any particular target leverage, but instead predicts that firms use first internal funds, followed by debt, and as a last resort, equity due to information asymmetries (Myers and Majluf, 1984; Myers, 1984). The market timing theory states that the capital structure of a firm is the cumulative outcome of attempts to time the equity market and gives no role to target leverage (see Baker and Wurgler, 2002; Huang and Ritter, 2009). Managerial inertia theory posits that stock returns passively derive the changes in market leverage ratios, rather than managers actively pursuing a target leverage (Welch, 2004).

Capital structure theories have different predictions regarding whether and how quickly firms converge to a target leverage. The pecking order, market timing, and inertia theories argue that there is no target capital structure and predict a speed of adjustment close to zero. The trade-off theory states that firms almost immediately reverse deviations from their target leverage, predicting a speed of adjustment close to one. Dynamic trade-off theory, on the other hand, considers market imperfections that cause delays in capital structure adjustments and predicts a speed of adjustment between zero and one. A higher estimated speed of adjustment supports more timely convergence to a target leverage. We add to this literature by exploiting machine learning prediction accuracy to reestimate the speed of adjustment. The new estimates allow us to better understand which theory of capital structure is supported by data.

Survey evidence indicates that target leverage is an important consideration for firms' capital structure decisions (Graham and Harvey, 2001). A long-standing question in the literature is how to measure unobservable leverage targets. Various studies have proposed different sets of factors associated with leverage. Firm size, R&D intensity, market-to-book, depreciation, tangibility, profitability, and median industry leverage are some commonly suggested determinants of target leverage (Harris and Raviv, 1991; Rajan and Zingales, 1995; Hovakimian et al., 2004; Flannery and Rangan, 2006; Öztekin and Flannery, 2012, among others). The possible omission of other relevant factors, however, can materially change inferences on existing factors and result in noisy target estimates that have critical implications for the validity of different capital structure theories (e.g., Flannery and Rangan, 2006; Lemmon et al., 2008; Frank and Goyal, 2009). We contribute to prior studies by training machine learning models to search among a large pool of

[1] Nonlinear patterns persist when we control for firm and time fixed effects (see Internet Appendix Figures A1, A2, and A3). This means that nonlinearities in the leverage models reflect both the time-series variation and cross-sectional differences, and fixed effects alone do not fully capture the hidden nonlinearities.

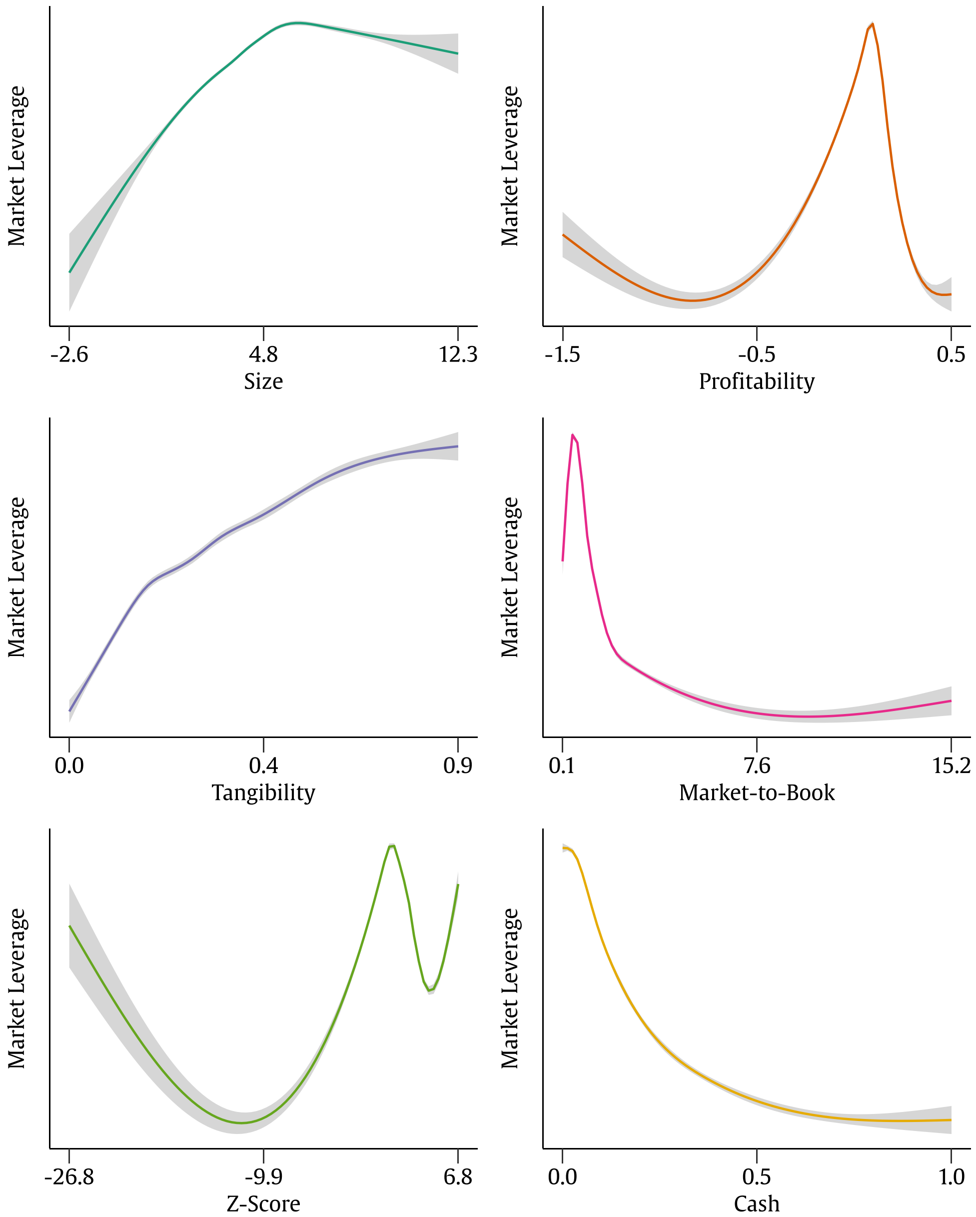

Fig. 1. Market leverage and firm characteristics.
This figure illustrates marginal relations between selected firm characteristics and corporate leverage using nonparametric spline-based smoothing functions. These estimates are obtained from our training sample over 1972–2000. The dependent variable is market leverage, defined as total debt scaled by the market value of assets (TDM). The description of selected variables is given in Internet Appendix Table A1.

potential determinants of capital structure while allowing for nonlinear patterns among these candidates.

A lack of consensus exists concerning the speed of adjustment as estimates considerably vary across different econometric methods. Some studies document slow convergence to the target (e.g., Fama and French, 2002; Welch, 2004; Kayhan and Titman, 2007; Huang and Ritter, 2009), and others document rapid convergence to the target (e.g., Leary and Roberts, 2005; Flannery and Rangan, 2006; Lemmon et al., 2008; Faulkender et al., 2012; Elsas and Florysiak, 2015). The literature attributes the lower speed of adjustment to noise in the estimation of target leverage and empirical models' failure to account for complex patterns in the data. Methods that allow for heterogeneous slopes and intercepts across firms typically generate faster adjustment speeds (Flannery and Rangan, 2006; Lemmon et al., 2008; Frank and Shen, 2019). While these studies rely on different econometric techniques, they all use linear panel data methods to estimate leverage targets. In contrast, our paper relies on machine learning models to estimate leverage targets and documents a significantly faster speed of adjustment compared to traditional linear models.

According to the dynamic leverage adjustment framework, firms converge to their target leverage faster when the costs of adjustment are low and/or the benefits of adjustment are high (Leary and Roberts, 2005; Strebulaev, 2007; Faulkender et al., 2012; Lockhart, 2014, among others). For example, Faulkender et al. (2012) document aggressive adjustment by firms with high absolute cash flows and high absolute leverage deviations. Korajczyk and Levy (2003) illustrate that unconstrained firms time their issue choice to coincide with periods of favorable macroeconomic conditions. DeAngelo et al. (2011) present a dynamic capital structure model in which firms sometimes issue transitory debt and deviate deliberately, but temporarily, from their leverage targets to fund investment. Çolak et al. (2018) illustrate that the estimated speed of leverage adjustment almost halves under high uncertainty due to elevated financial intermediation costs. While there is empirical support for different variables affecting the adjustment speed, existing studies test the featured factors in isolation. We contribute to this literature by simultaneously evaluating a large set of adjustment speed determinants and assessing their relative importance in a machine learning framework.

Our study speaks to the expansive literature on the determinants of capital structure. Frank and Goyal (2009) apply a linear model selection method using Akaike information criterion (AIC) and Bayesian information criterion (BIC) to identify the determinants of corporate leverage. From a large pool of potential leverage factors for U.S. firms, their work finds support for industry median leverage, market-to-book, profitability, tangibility, firm size, and expected inflation. Öztekin (2015) utilizes a sample of international firms and linear models and finds similar results. Neither paper, however, considers nonlinearities that are present in the relation between corporate leverage and its determinants. Machine learning models can potentially resolve this issue as they, by design, model complex and nonlinear patterns more accurately. Exploiting these models, we document substantial gains in the accuracy of predicting target leverage and find a more extensive set of capital structure determinants. Additionally, two influential studies, Lemmon et al. (2008) and DeAngelo and Roll (2015), examine the stability of corporate leverage, whereas our study examines the stability of leverage determinants.

Our paper also extends empirical studies investigating whether managers strive to converge to their target leverage (e.g., Welch, 2004; Flannery and Rangan, 2006; Lemmon et al., 2008; Öztekin and Flannery, 2012). Except for Frank and Shen (2019), previous work does not directly assess the impact of precision in estimating leverage targets on the adjustment speed and prediction of firms' financing behavior. Leverage targets estimated by linear methods exhibit slow adjustment and incorrect predictions of observed firms' financing actions. Conversely, our machine learning methods minimize the measurement error in predicting leverage, leading to a sharp increase in the estimated speed of leverage adjustment and to more theoretically-consistent predictions of firms' financing patterns.

Lastly, our work adds to a new and growing literature applying machine learning estimators to different research areas in finance. For instance, Gu et al. (2020) use machine learning methods to predict stock returns as well as the most informative predictors of the returns. Erel et al. (2021) exploit machine learning's superior predictive performance to determine which potential director would be the best fit for a given firm. Li et al. (2021) use machine learning techniques including a NNET to calculate the association between words appearing in earnings call transcripts to quantify corporate culture. Bianchi et al. (2020) and Bali et al. (2020) employ machine learning to reconcile the debate on the presence of predictable variation in bond returns. Bubna et al. (2020) use unsupervised machine learning tools to study the characteristics of venture capitalists that enter syndicate partnerships with each other. To the best of our knowledge, our paper is the first to apply machine learning algorithms to a large pool of variables to identify the most important predictors of leverage and estimate leverage targets as well as the speed of adjustment to those targets.

3. Empirical methodology

In this section, we describe the empirical models employed in this study for predicting capital structure dynamics. In the first step, we use machine learning to improve the prediction of target leverage and compute firms' deviations from their target leverage. In the second step, we use a partial adjustment framework to estimate the speed of adjustment to target leverage.

The basic regression problem in predicting target capital structure is to estimate a function $g(\mathbf{x}_{i,t}) = E(y_{i,t+1}|\mathbf{x}_{i,t})$ where

$$y_{i,t+1} = g\left(\mathbf{x}_{i,t}\right) + \varepsilon_{i,t+1} \tag{1}$$

is the ith firm's target leverage ratio in year t+1 and $\varepsilon_{i,t+1}$ is a random error component. The regression function $E(y_{i,t+1}|\mathbf{x}_{i,t})$ is the conditional expectation of $y_{i,t+1}$ conditioned on the vector of covariates described in the next section.

Our goal is to estimate the function $g(\cdot)$ using several methods ranging from standard multiple linear regression to highly nonlinear, discontinuous methods such as a random forest. The first class of models presented in Section 3.1 are variants of a standard linear model. Note, however, that they are not necessarily linear in the covariates. In fact, the generalized additive model described in Section 3.1.3 can be used to estimate highly nonlinear relations between the covariates and the dependent variable using splines or any suitable nonparametric function estimation tool. The second class of models in Section 3.2, referred to here as machine learning prediction functions, allows for nonlinear and discontinuous relations between y_i and the associated covariates. Further, these machine learning models may capture complex interaction effects that are present in the covariates. See Hastie et al. (2009) and Kuhn and Johnson (2013) for additional information related to the models considered here.

3.1. Linear models

3.1.1. Multiple regression model

Our baseline model is the standard, multiple regression model of the form

$$g(\mathbf{x}_{i,t};\boldsymbol{\beta}) = \mathbf{x}_{i,t}^{'}\boldsymbol{\beta}. \tag{2}$$

The regression parameters $\boldsymbol{\beta}$ estimated using ordinary least squares (OLS) are defined by

$$\widehat{\boldsymbol{\beta}}^{ols} = \underset{\beta}{\operatorname{argmin}} \parallel \mathbf{y} - \mathbf{x}\boldsymbol{\beta} \parallel_2^2, \tag{3}$$

where $\|\mathbf{a} - \mathbf{b}\|_2$ is the distance (Euclidean) between the vectors $\mathbf{a}$ and $\mathbf{b}$, i.e., the l_2-norm.

3.1.2. Lasso

There are two drawbacks to using a standard multiple regression model in our study: (1) overfitting and (2) potential multicollinearities given a large set of predictor variables. To mitigate these concerns, we utilize the least absolute shrinkage and selection operator model, also known as the LASSO (Tibshirani, 1996). LASSO is an effective method for shrinking parameters associated with insignificant covariates to zero. Thus, LASSO acts as both a shrinkage *and* a model selection tool, and the result is a sparse version of the standard multiple regression model defined in Eq. (2). As a result, LASSO should be able to select the appropriate variables in a more general capital structure model, by including only a subset of the original predictor variables that are important for predicting target leverage.

The parameter estimates in the LASSO model for a given value of λ are given by

$$\widehat{\boldsymbol{\beta}}_{\lambda}^{lasso} = \underset{\beta}{\operatorname{argmin}} \parallel \mathbf{y} - \mathbf{x}\boldsymbol{\beta} \parallel_2^2 + \lambda \parallel \boldsymbol{\beta} \parallel_1 \tag{4}$$

for some $\lambda > 0$. The effect of λ and the l_1-penalty term ($\lambda \parallel \boldsymbol{\beta} \|_1 = \lambda \sum_{j=1}^{p} | \beta_j |$) is straightforward. If $\lambda = 0$, then $\widehat{\boldsymbol{\beta}}^{ols} = \widehat{\boldsymbol{\beta}}^{lasso}$ and as $\lambda \to \infty$, the penalty term forces the coefficients to zero. Data-driven choices of the tuning parameter λ are usually obtained using some form of cross-validation, as we discuss in Section 3.3.

3.1.3. Generalized additive model

OLS and LASSO models are not well equipped to adequately model nonlinear relations between the covariates and the dependent variable without pre-specified manual intervention. A generalized additive model, or GAM, is a family of nonparametric models that can handle the nonlinearities that may be present (James et al., 2013). In the case of GAM, the regression function $g(\cdot)$ defined in Eq. (1) can be written as

$$g(\mathbf{x}_{i,t}) = \beta_0 + \sum_{j=1}^{p} f_j(x_{ij,t}) \tag{5}$$

for smooth, nonlinear functions f_j. Spline functions are often used in estimating the functions f_j, however, almost any standard nonparametric function estimator can be used, e.g., local polynomial regressions estimators. One of the main advantages of using GAM as opposed to a classic machine learning algorithm for modeling nonlinear relations is that the additive, functional effects of the covariates x_j on y are readily available and are often interpretable.

3.2. Machine learning models

Given that prediction and identification of variable importance are among our central goals in this paper and nonlinear models might be more apt in the capital structure domain, we investigate the utility of several machine learning (ML) models in our study. ML algorithms are particularly useful for modeling nonlinear relations among dependent and independent variables as well as capturing hidden interactions in these variables. And, although nonlinear relations and interaction effects can be modeled using the methods described in Section 3.1, they must be noted a priori. In contrast, the ML methods described here are fully nonparametric and sufficiently flexible to capture these complex structures without prior intervention.

The specific ML models utilized here and described in Sections 3.2.1, 3.2.2, and 3.2.3 are a random forest (Breiman, 2001), a gradient boosting machine (Friedman, 2001), and a feed-forward neural network with a single hidden layer (Venables and Ripley, 1997), respectively. Additional algorithmic details related to these methods can be found in Kuhn and Johnson (2013) and James et al. (2013). We adopt their terminology and refer to independent variables as "features" and the collection of feature vectors as the "feature space."

3.2.1. Random forests

To motivate a random forest (RF), we first describe two key aspects of a simple regression tree. First, suppose we split the feature space $\mathscr{X}$ into J unique, nonoverlapping regions: $R_1, R_2, \ldots, R_J$. The predicted value of y for any value within R_j is the average overall response values in R_j, or more formally:

$$\widehat{g}^{rf}(\mathbf{x}) = \sum_{j=1}^{J} \overline{y}_j I_{\{\mathbf{x} \in R_j\}}, \tag{6}$$

where $I_{\{\mathbf{x}\in R_j\}}$ is an indicator function that equals one if $\mathbf{x}$ is in R_j and zero otherwise. The residual sum of squares (RSS) can then be computed in a clear manner. The predictions made from growing a single tree are notorious for exhibiting high variance. That is, they may change substantially (highly variable) from sample to sample. The method of bootstrap aggregation, or bagging, is often employed to alleviate this potential problem. In a regression context, bagging starts by taking a bootstrap (see Efron and Tibshirani, 1994) sample and growing a regression tree on this sample. As is usually the case in utilizing a bootstrap methodology, many bootstrap samples are generated, say B, and trees are grown on each sample. This results in an ensemble (or forest) of trees from which to make predictions. The bagged estimate at $\mathbf{x}$ is the average estimate over all trees:

$$\widehat{g}^{bag}(\mathbf{x}) = \frac{1}{B}\sum_{b=1}^{B}\widehat{g}^{b}(\mathbf{x}), \tag{7}$$

where $\widehat{g}^{b}$ is the estimator defined in Eq. (6) on the bth bootstrap sample. The effect of averaging over this ensemble is to reduce the variation in the final estimate (Breiman, 1996).

The variance of the bagged estimate in Eq. (7), however, can still be large due to averaging over potentially highly correlated estimates. When using the entire collection of features, each tree tends to be correlated to the trees grown from the other bootstrap samples, particularly when several features tend to dominate the trees. The random forest has been introduced as a method to "decorrelate" the bootstrap-grown trees (Breiman, 2001).

Each tree in a random forest is grown on a bootstrap sample selected from the original data set. The random forest differs from a bagged approach in that only a random subset, say $m < p$, of the features are considered as candidates for partitioning the feature space during the growing process. This attenuates the effect of dominant features and promotes features that tend to be masked, resulting in a collection of trees that are less correlated than their bagged analogs. The final estimate is still constructed as defined in Eq. (7).

3.2.2. Gradient boosting regression trees

Compared to bagging and RF algorithms that use bootstrapping, the GBM sequentially grows trees by updating the data used to grow a tree after each tree is fit. The algorithm starts by fitting a tree using the original data set. Subsequent trees are grown using the fitted residuals, updated after each fit. The final estimate of $\widehat{g}_{gbm}(\mathbf{x})$ is the weighted sum of the individual estimates from each tree. The weighting is controlled by a parameter λ that determines how fast the model "learns" and is usually found via cross-validation.

3.2.3. Neural networks

Our final choice of machine learning algorithms is a single-layer, feed-forward neural network (1-FFNL) (see Venables and Ripley, 1997; Gu et al., 2020). In this framework, the predictor variables are fed into a hidden layer that transforms the predictors in a (possibly) nonlinear and interactive manner. The hidden layer is then combined to estimate the response variable.

3.3. Model tuning

Machine learning models usually have one or more parameters not directly estimable from the data, often referred to as tuning parameters. In this situation, a candidate set of tuning parameters can be selected a priori to evaluate their effectiveness (see Kuhn and Johnson, 2013, for tuning process). For each candidate parameter, the model is fit to a subset of the data and then evaluated on its performance at predicting new observations in which the user knows the true response variables. The final tuning parameters are chosen based on pre-specified optimality criteria. Internet Appendix Table A2 reports the tuning parameters used in our final models.

To evaluate a parameter's effectiveness on model performance, we use subsampling from the original data set. We divide our data into training sets (1972–2000) and testing sets (2001–2018). Within the training set, we create 10 subsamples, or folds, in the ML parlance, upon which we tune and validate each model's performance on a validation set. Internet Appendix Fig. A4 illustrates the tuning/validation process used in our analysis. The figure's horizontal axis represents every year contained in our training set. Each row in this figure highlights a fold, or subsample, and illustrates the data by year that are used to tune our models (blue squares) followed by the validation year (red squares). The data in red squares are often referred to as a validation sample. This is a standard 10-fold cross-validation method across time and should ensure a robust set of tuning parameters (see Kuhn and Johnson, 2013).

The process evolves by training each model using its respective candidate tuning parameters on the individual 10 subsamples. The tuning parameters' performance is evaluated based on how well the trained model predicts leverage in the validation sets averaged across the subsamples. For example, a researcher might choose the tuning parameters that yield the lowest mean squared forecast error (MSFE) on average.

3.4. Model fitting

Once the appropriate set of tuning parameters is found for its respective model, we begin by fitting each candidate model to the entire training set from 1972 to 2000 using those tuning parameters. We use this model to make out-of-sample predictions in 2001. The model is updated in each subsequent year, e.g., using data from 1972 to 2001, to make predictions on the following year, 2002. We repeat this process through 2018, however, the tuning parameters remain the same.

4. Sample selection and summary statistics

4.1. Variable selection and data sources

The empirical literature that tests the determinants of capital structure applies different sets of potential variables implied by each

Table 1
Summary statistics for the aggregate sample.

	Observations	Q1	Mean	Median	Q3	Std.
Panel A: Firm-Level Characteristics:						
Leverage Measures:						
TDM	202,061	0.033	0.274	0.205	0.455	0.261
TDA	231,648	0.052	0.244	0.202	0.374	0.224
LDM	202,061	0.007	0.205	0.130	0.344	0.221
LDA	232,130	0.011	0.180	0.122	0.289	0.195
Profitability:						
Profit	226,831	0.020	0.051	0.095	0.162	0.240
Firm Size:						
Assets	233,225	3.590	5.193	5.101	6.704	2.245
Mature	233,225	0.000	0.747	1.000	1.000	0.434
Growth:						
Mktbk	202,061	0.704	1.603	1.021	1.716	1.867
ChgAsset	227,710	−0.024	0.126	0.072	0.199	0.445
Capex	216,700	0.014	0.062	0.039	0.079	0.075
Nature of Assets:						
Tang	227,820	0.056	0.264	0.192	0.400	0.247
R&D	229,170	0.000	0.213	0.000	0.023	1.452
Unique	225,021	0.000	0.235	0.000	0.000	0.424
SGA	190,409	0.141	0.404	0.236	0.374	0.886
Cash	232,420	0.024	0.159	0.070	0.205	0.207
Taxes:						
TaxRate	233,225	0.350	0.380	0.350	0.460	0.058
Depr	215,060	0.022	0.046	0.037	0.057	0.043
InvTaxCr	209,254	0.000	0.001	0.000	0.000	0.004
Risk:						
StockVar	230,480	0.000	0.002	0.001	0.002	0.009
Z-Score	187,768	0.638	1.017	1.793	2.764	3.793
Supply-Side Factors:						
Rating	233,225	0.000	0.110	0.000	0.000	0.312
Stock Market Conditions:						
StockRet	217,712	−0.241	0.152	0.048	0.358	0.835
CrspRet	217,712	−0.008	0.115	0.130	0.253	0.173
Panel B: Industry-Level Characteristics:						
Industry:						
IndustLev	216,178	0.069	0.242	0.210	0.381	0.197
IndustGr	233,035	0.029	0.086	0.076	0.121	0.163
Regultd	225,021	0.000	0.034	0.000	0.000	0.181
Panel C: Macro-Level Characteristics:						
Debt Market Conditions:						
TermSprd	233,225	−0.045	0.019	0.008	0.075	0.079
Macroeconomic Conditions:						
Inflation	233,225	0.024	0.041	0.033	0.048	0.023
MacroProf	233,225	−0.060	0.058	0.090	0.163	0.131
MacroGr	233,225	0.019	0.028	0.031	0.041	0.019
Panel D: Other Characteristics:						
Refinancing Proxies*:*						
NetDebt	200,540	−0.008	−0.041	0.000	0.019	0.199
NetPay	206,945	−0.016	0.012	0.000	0.027	0.103

The sample includes U.S. firms listed on NYSE, AMEX, or NASDAQ with CRSP share codes of 10 and 11, which are covered by CRSP and Compustat between 1972 and 2018. We exclude utility firms (SIC codes 4900–4999) and financial firms (SIC codes 6000–6999) as well as firms for which total assets are either missing or negative. The table reports summary statistics for the sample including the number of observations, first quartile (Q1), mean, median, third quartile (Q3), and standard deviations (Std.). All ratios are winsorized at the 0.5th and 99.5th percentiles of their empirical distributions. The description of variables is provided in Internet Appendix Table A1.

theory. Whether these factors affect corporate leverage positively or negatively depends on the assumptions behind the underlying theory. The potential determinants of corporate leverage roughly belong in three categories: (1) firm-level characteristics, including proxies for size, profitability, growth opportunities, bankruptcy risk, and nature of assets; (2) industry-level characteristics, including proxies for industry leverage, industry growth, and whether the firm operates in a regulated industry; and (3) macro-level characteristics, including proxies for macroeconomic conditions, stock market conditions, debt market conditions, tax policies, and accessibility to the debt market. We collect data from multiple sources to construct different proxies within each category.

Our primary sample includes U.S. firms listed on the NYSE, AMEX, or NASDAQ with CRSP share codes of 10 and 11, which are covered by CRSP and Compustat between 1972 and 2018. We exclude firms for which sales and total assets are either missing or negative as well as utility firms (SIC codes 4900–4949) and financial firms (SIC codes 6000–6999). Following Hou et al. (2015), we use CRSP SIC codes whenever Compustat SIC codes are not available. Data on the expected inflation rate are from the Livingston Survey, conducted by the Federal Reserve Bank of Philadelphia. Growth in corporate profits and GDP are from the Federal Reserve Bank of St. Louis Economic Data (FRED). Internet Appendix Table A1 describes the variables used in this paper and their sources. We also refer the reader to Frank and Goyal (2009) and Kayo and Kimura (2011), who provide detailed discussions on these proxies and the theories that predict their signs. At times, some of the data possess outlier problems. To cope with extreme values and mitigate their effect, we winsorize all ratios at the top and bottom 0.5 percentiles of their underlying distributions in the entire sample.

4.2. *Descriptive statistics*

Our sample closely follows Frank and Goyal (2009) with a few changes. We extend their sample by 15 years and add additional leverage factors including Altman's (1968) *Z*-Score to proxy for the firm's risk and cash balances to proxy for the firm's liquidity. We exclude net operating loss carryforward (NOLCF) from our sample due to a high percentage of missing observations. Moreover, our sample begins in 1972, not 1950, as 1972 is the earliest year for which we have observations for the entire set of potential determinants of capital structure. After keeping the nonmissing observations across all the control variables and dependent variables, we are left with 128,417 firm-year observations. We lag all the right-hand-side variables by one year in our empirical models to mitigate potential endogeneity issues and to ensure that all variables are publicly known at the time of prediction.

Table 1 reports the summary statistics. An average firm in our sample has a market leverage (TDM) ratio of 27.4% and enjoys a 5.1% profit rate, where profit is operating income before depreciation scaled by total assets. Capital expenditures account for 6.2% of total assets, R&D expenditures are 21.3% of total sales, and cash and equivalents account for 15.9% of total assets. Investing in an average firm in our sample yields an annual return of 15.2%. The empirical capital structure literature uses alternative measures of debt ratio, and these proxies differ in the choice of scaling factor and whether total debt or long-term debt is considered. We report the majority of our results using the most common measure of capital structure, market leverage, defined as the ratio of total debt to the market value of assets (TDM). Our conclusions using alternative measures of leverage are similar.

5. Machine learning and leverage

5.1. *Predictive performance*

For the main results, we initially consider TDM as the dependent variable and the 26 variables listed in Panels A, B, and C of Table 1 (starting with "Profit" and ending with "MacroGr") as the independent variables. We use 1972–2000 as the training period to tune our machine learning models and test the out-of-sample performance of these models over the 2001–2018 period. To evaluate the predictive performance of each machine learning model at time $t+1$ in the testing period, we estimate two closely related metrics: mean squared forecast error (MSFE) and the out-of-sample R-squared (R_{OS}^2), defined as

$$MSFE = \frac{1}{N}\sum_{i=1}^{N}\left(y_{i,t+1} - \hat{y}_{i,t+1}\right)^2, \text{and} \tag{8}$$

$$R_{OS}^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_{i,t+1} - \hat{y}_{i,t+1}\right)^2}{\sum_{i=1}^{N}\left(y_{i,t+1} - \overline{y}_{t+1}\right)^2}, \tag{9}$$

where N is the number of observations at time t+1, $y_{i,t+1}$ is the ith actual response at time t+1, $\hat{y}_{i,t+1}$ is the ith fitted value based on the selected model, and $\overline{y}_{t+1}$ is the average of response at time t+1.

Fig. 2 presents R_{OS}^2 statistics for our machine learning models as well as the linear (LM) and LASSO models over the test period (2001–2018). The linear model generates R_{OS}^2 statistics ranging from approximately 28.0% in 2004 to 44.0% in 2015. LASSO's performance is roughly close to the LM, albeit with less variation, ranging from 30.6% to 39.4%. Overall, RF has the highest consistent forecasting power, with R_{OS}^2 ranging between 45.6% in 2009 to 58.7% in 2012. The relatively low predictability in 2009 is likely driven by the financial crisis and its recovery, which drove large changes in investment, borrowing, and market values. Nonetheless, predictability over 45% still indicates relatively robust predictive performance.

Predictability for GBM nearly matches RF's R_{OS}^2 throughout the sample and reaches 55.3% by 2018. GBM's R_{OS}^2 also takes a modest

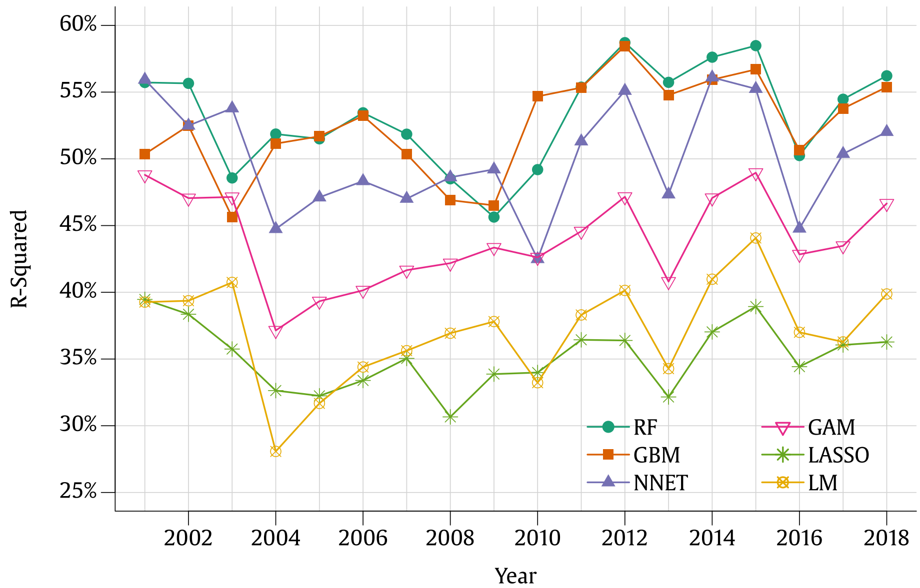

Fig. 2. Out-of-sample performance for market leverage.
This figure plots the out-of-sample R-squared, R^2_{OS}, for different machine learning models and a linear model (LM). The out-of-sample period is 2001–2018, and the dependent variable is market leverage, defined as total debt scaled by the market value of assets (TDM). Machine learning models used in the analysis are random forests (RFs), gradient boosting machines (GBMs), neural networks (NNETs), general additive models (GAMs), and least absolute shrinkage and selection operator (LASSO) models.

hit during the financial crisis, dipping to approximately 46%. NNET R^2_{OS} ranges from 42.4% in 2010 to 56.1% in 2014, while GAM R^2_{OS} ranges from 37.1% in 2004 to 48.9% in 2015. Overall, results indicate that all machine learning models improve on the performance of the linear OLS and LASSO models, and R^2_{OS} remains relatively stable from 2001 to 2018. The RF model emerges as the best predictor and generates substantial forecasting gains over conventional linear models. The difference between the R^2_{OS} of the RF and LM models ranges from 7.8% in 2009 to 23.7% in 2004.

Goyal and Welch (2003) introduce a method for evaluating the out-of-sample forecasting ability of predictive regressions. They graphically illustrate the difference over the out-of-sample period between the forecast accuracy of a benchmark and alternative model. This contrasts to statistics such as MSFE or R^2_{OS}, which are point estimates and hence mask the stability of the predictive performance over time. To evaluate predictability over the past 18 years, Internet Appendix Fig. A5 plots the cumulative percentage difference between the MSFE of LM and our machine learning models over the test period (2001–2018). This figure illustrates performance relative to the LM, in contrast to the prior figure, which illustrates out-of-sample predictability relative to the historical average constant. An upward slope indicates that the errors for the benchmark model exceed those of machine learning models. The steady positive slopes of RF, GBM, NNET, and GAM imply that these models consistently outperform the linear model. All four methods have a positive slope every year, indicating that their forecasts generate lower out-of-sample errors every year compared to the LM. On the other hand, LASSO's relatively flat and then downward slope suggests that this method's variable selection might be too parsimonious since it leads to similar or even larger errors than the LM.

Table 2 presents the out-of-sample predictability for the last year of our sample (2018). The LM generates an R^2_{OS} of 39.8%, and it outperforms LASSO's performance by 3.6%. The two best performers are RF and GBM with R^2_{OS} of 56.2% and 55.3%, respectively. Compared to the linear model, RF and GBM R^2_{OS} represent a large increase in out-of-sample predictability. The performance compares favorably with Frank and Goyal (2009), who use backward stepwise regressions; their in-sample R^2 that minimizes the AIC/BIC criteria is 26.6%. Typically, R^2_{OS} statistics are lower than their in-sample counterpart due to less overfitting. Further, low R^2_{OS} statistics may indicate model failure due to parameter instability. However, since predictability is relatively high, and the R^2_{OS} is stable over time, it is unlikely that the machine learning models of corporate leverage lead to predictive failure arising from parameter instability.

Table 2 also presents the root mean squared error (RMSE). LM and LASSO have RMSE of 0.190 and 0.196, respectively, while GBM and RF's RMSE values are 0.164 and 0.162, respectively, representing sizeable reductions of 13.7% and 14.7%. To evaluate the significance of their predictive accuracy, we use a panel Diebold Mariano test (Ductor et al., 2014). The null is equal predictability, and the alternative is that machine learning models contain useful information beyond the LM. Results document that RF, GBM, NNET, and

Table 2
Out-of-sample performance of machine learning versus linear models.

Model	R^2_{OS}	RMSE	RMSE Differential (%)
LM	0.398	0.190	0.000
LASSO	0.362	0.196	+3.158***
GAM	0.466	0.179	−5.790***
NNET	0.520	0.170	−10.526***
GBM	0.553	0.164	−13.684***
RF	0.562	0.162	−14.737***

This table reports the out-of-sample R-squared, R^2_{OS}, and the root mean squared error, RMSE, for linear models (LM), least absolute shrinkage and selection operator (LASSO) models, generalized additive models (GAM), neural networks (NNET), gradient boosting machines (GBM), and random forests (RF). The estimates are reported for the last year of the sample (2018). The dependent variable is market leverage, defined as total debt scaled by the market value of assets (TDM), and the description of control variables is given in Internet Appendix Table A1. The last column reports the percentage difference between the corresponding model's RMSE and the LM as the benchmark. Symbols ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively, for the panel Diebold Mariano test.

GAM possess significantly lower RMSE than the benchmark LM. Hence, nonlinear machine learning models generate statistically significant increases in forecast accuracy of leverage targets.

5.2. Variable importance

We now turn our attention to identifying covariates that have substantial power in explaining the variation in corporate leverage. We illustrate each variable's contribution to predicting TDM with variable importance plots for four of our machine learning models estimated for the first year of the out-of-sample period (2001). GAM is not presented because it does not measure variable importance. Additionally, to examine whether the determinants of leverage have evolved over time, we present variable importance for the last year of the out-of-sample period (2018) to evaluate potential changes in importance. In the context of machine learning, Garson (1991) and Goh (1995) identify the relative importance of explanatory variables by deconstructing the model weights. The intuition is that the relative importance or strength of association of a particular covariate can be determined by identifying all weighted connections between the nodes of interest. Variables with relatively high importance are drivers of the forecasts and their estimates have a larger impact on leverage than values with low importance. For comparison, each model's variable importance measures are normalized so that the most important variable equals 100. All other variables are then scaled relative to this maximum.

We focus on the RF model since it consistently exhibits the greatest predictive power in our study. Archer and Kimes (2008) demonstrate that RF outperforms other machine learning methods in determining the covariates that best contribute to the predictive structure. Similarly, Strobl et al. (2008) find that tree-based methods such as RF excel at identifying relevant predictor variables even in high-dimensional settings involving complex interactions. Moreover, RF models use a feature bagging method, which has the advantage of decreasing the correlation between each decision tree and thus, on average, increasing its predictive accuracy (Breiman, 2001).

Turning to each control variable's contribution to explaining corporate leverage, Fig. 3 presents the variable importance fitted to the training data (1972–2000). Panel A documents that RF picks market-to-book as the most important predictive variable and this value is then normalized to 100. Industry median leverage, cash, *Z*-Score, profitability, stock returns, and firm size are also important, with weights of at least 20% relative to the most important variable. This implies that RF identifies more variables as important than the work by Rajan and Zingales (1995); their work uses market-to-book, profitability, tangibility, and size as drivers of corporate leverage. We find that industry conditions, stock market conditions, and risk factors are also relevant determinants of corporate leverage. Altogether, the RF model has 20 variables with weights exceeding 5% relative to the most important variable.

Similarly, Panel B of Fig. 3 shows that GBM and RF generate relatively similar variable importance results. GBM weights market-to-book as the most important factor, although industry median leverage, cash, Z-Score, and profitability also have weights higher than 20%. The key difference with RF is that the GBM weights tend to drop off more rapidly. The equitability of weights is even more relevant for NNET. The NNET variable importance plot shows that R&D expenditure is the most important leverage determinant, although profitability, cash, stock variance, and market-to-book also have at least 50% weights. Further, another seven variables have weights exceeding 20%, including depreciation and amortization; growth in GDP; expected inflation; debt ratings; regulated industry; investment tax credit; and selling, general, and administrative expenses (SGA). Hence, NNET more equally assigns variable importance than RF, GBM, and LASSO.[2] The LASSO procedure, which is inherently a linear model, weights industry leverage as the most

[2] We limit our study of NNET to one layer, due to our focus on the determinants of capital structure, as multi layer NNET does not readily provide variable importance information. It is also worth noting that the black-box nature of neural networks can make the interpretation of the individual variable's contribution to prediction difficult. To rectify this problem, we employ the R Software package NeuralNetTools' Garson method (Garson, 1991) to estimate variable importance. See Beck (2018) and R Core Team (2019) for additional information related to the NeuralNetTools package and the R Software, respectively.

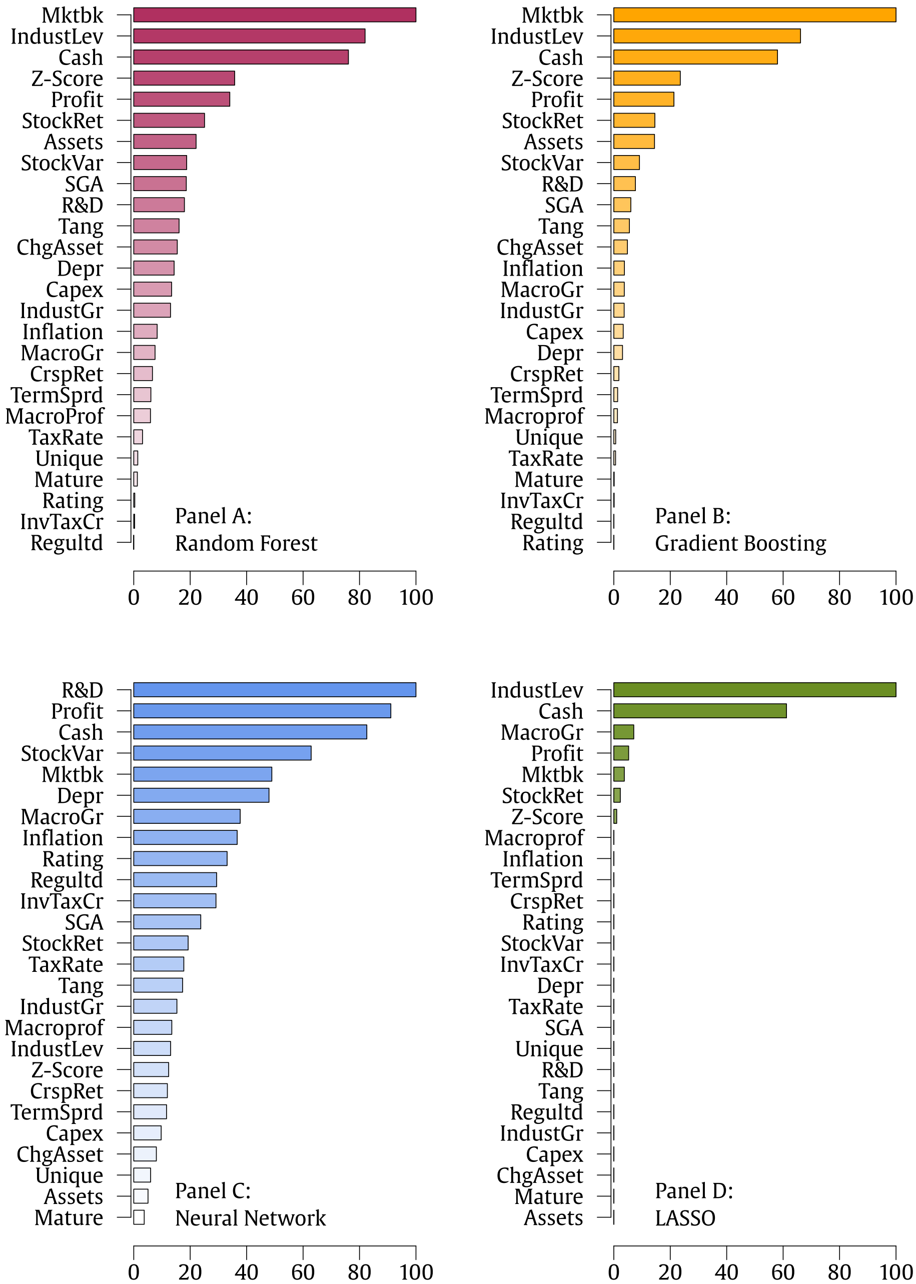

Fig. 3. Initial variable importance obtained from different machine learning models.
This figure plots the importance of explanatory variables for predicting target leverage using random forest (RF), neural network (NNET), gradient boosting machine (GBM), and least absolute shrinkage and selection operator (LASSO) estimators. The variable with the highest importance is normalized to 100. These estimates are obtained from our training sample over 1972–2000 and used to calculate variable importance for 2001. The dependent variable is market leverage, defined as total debt scaled by the market value of assets (TDM). The description of control variables is given in Internet Appendix Table A1.

important, followed by cash, with at least 20% weight. Other non-zero-weight variables picked by LASSO include growth in GDP, profitability, market-to-book, stock returns, and *Z*-Score. The other variables receive a weight of zero.

Graham and Leary (2011) provide evidence that several of the suggested leverage determinants affect the debt ratio in a nonlinear fashion; yet, few studies explicitly model these nonlinearities when predicting target leverage or quantifying variable importance. Our study takes a step to address these issues and shows that incorporating nonlinearities into capital structure models (here through machine learning estimators) changes our understanding of what factors are reliably important. While linear models such as LASSO select only a few factors as important drivers of leverage, machine learning models draw predictive information from a much broader set of covariates. It is therefore likely that prior studies are unable to detect additional leverage factors because they assume that the relation between the potential covariates and target leverage is linear, whereas theory suggests, and our results confirm, that it is not.[3]

To evaluate the evolution of variable importance over time, we use two methods. In the first method, we plot variable importance in Internet Appendix Fig. A6 for data fitted from 1972 to 2017 to predict TDM in 2018. We then compare these plots to Fig. 3. In Panel A, the RF model displays only modest reordering of the top three variables. For instance, industry median leverage becomes the most important factor over the full sample and market-to-book moves from first to third place. Cash, Z-Score, profitability, stock returns, firm size, SGA, and tangibility still emerge as important determinants of leverage 17 years later, with weights of at least 20%. GBM, plotted in Panel B, also displays remarkable variable stability; it is similar to 17 years earlier, with 12 variables weighted over 5%. Additionally, GBM ranks fewer variables as "very" important relative to RF; e.g., only four variables possess an importance rank of over 20%, and a dozen variables have a relative importance between 5% and 20%. Turning to Panel C, NNET displays modest instability in the ranking of some variables. For example, profitability is the second important variable originally, but then falls to seventh by the end of the sample period. Other variables do not change substantially in importance; stock variance, cash, R&D expenditure, depreciation and amortization, expected inflation, and market-to-book are the top six variables in Internet Appendix Fig. A6, and these variables are among the top eight in Fig. 3. Finally, in Panel D, LASSO's ranks in 2018 remain identical to those at the beginning of the testing period with only a few exceptions. Industry leverage is ranked as the most important, followed by cash, profitability, market-to-book, stock returns, Z-Score, and tangibility with nonzero weights. All other variables possess zero weights. Remarkably, the top 10 variables in all panels of Fig. 3 are also ranked in the top 10 in the panels of Internet Appendix Fig. A6. Again, there is only moderate reordering within the top 10 and 20. Industry median leverage is now ranked as the most important factor by three out of four models. Compared to the other three methods, LASSO assigns most variables with low weights, and hence leverage is determined by only a few variables. However, LASSO's relatively low predictability compared to RF and GBM, documented in Table 2, suggests that leverage is determined by a wide array of variables and nonlinearities.

In the second method, we illustrate the stability of variable importance in Fig. 4 and focus on RF, our top-performing model. For clarity and conciseness, we plot variables with importance over 20%. Results show that market-to-book has the highest predictive power from 2001 to 2012, and then its relative importance declines as both industry median leverage and cash climb in their importance. This change implies that managers over time tend to push the firm debt ratio closer to the debt ratio of peer firms that are in the same industry. The figure conveys another central message: variable importance is relatively stable over time. The top three performers are the same for the full out-of-sample period, and the other four variables, Z Score, profitability, stock returns, and firm size, possess relatively flat slopes, indicating that their relative contribution remains similar over the out-of-sample period.

How do our variable importance results compare to other related studies? Frank and Goyal (2009) employ a model selection approach using Akaike information criterion (AIC) and Bayesian information criterion (BIC) to identify a set of reliable corporate leverage determinants. They find six core factors that explain 26.6% of the variation in leverage: industry median leverage, tangibility, market-to-book, profitability, firm size, and expected inflation. Results for RF also highlight the importance of these variables. However, RF additionally weights cash, *Z*-Score, stock returns, and selling, general, and administrative expenses with weights exceeding 20% in Internet Appendix Fig. A6 and Fig. 4. GBM further highlights the relevance of cash, Z-Score, and stock returns with high weights. Work by Kayo and Kimura (2011) emphasizes fewer variables, including growth opportunities, profitability, distance from bankruptcy, size, and tangibility.

To mitigate the potential absorption of control variables' explanatory power, and to ensure that our results are comparable to related studies (e.g., Rajan and Zingales, 1995; Frank and Goyal, 2009), we do not include firm fixed effects in the analyses above. For completeness, Internet Appendix Table A3 compares the predictive performance of the machine learning methods to those of conventional models after controlling for firm fixed effects. Our main conclusions remain unchanged. RF and GBM models produce R^2_{OS} statistics that are substantially higher than the statistics produced by other models.[4] We also find that the relative ranking of the leverage determinants is not materially affected.

5.3. Alternative measures of leverage

Although many studies advocate market-based leverage as a more managerially relevant measure of corporate leverage due to its forward-looking nature (e.g., Welch, 2004), others support the use of book-based corporate leverage measures. Since a firm's book values exclude growth opportunities, these values primarily reflect tangible assets, which can be used as debt collateral. Therefore, managers are more likely to set their debt issuance policies based on book leverage. A survey by Graham and Harvey (2001) provides support for this view. Barclay et al. (1995) find that practitioners typically use book leverage as the proxy for corporate leverage. This section investigates how predictability and variable importance differ for our three other measures of leverage: LDM, TDA, and LDA.

[3] Tax-based models suggest that the relation between optimal debt level and business risk is U-shaped under corporate and personal taxation (e.g., Kale et al., 1991). This is because the tax liability is an option portfolio that is long in the corporate tax option and short in the personal tax option. The net effect of a change in business risk on the optimal debt level depends on the relative magnitudes of the marginal changes in the values of these two options.

[4] For Internet Appendix Table A3, the inclusion of firm fixed effects in the machine learning models is conducted by demeaning the data across firms. As a result, the R^2_{OS} values in this table show the *remaining* explanatory power of the control variables *after* removing the explanatory power of firm fixed effects. The estimates highlight that the predictive performance of RF still dominates the standard linear model as the R^2_{OS} of RF is about 9.3% higher than that of the standard linear model.

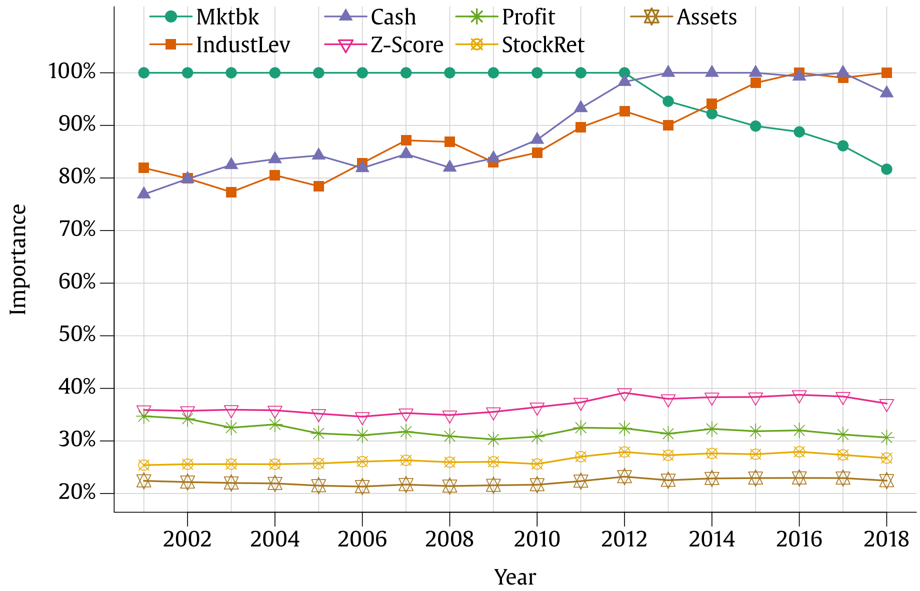

Fig. 4. The stability of variable importance over time for market leverage.
This figure plots all variables whose importance exceeds 20% over the testing period, 2001–2018, for predicting target leverage using random forests (RFs). The variable with the highest importance is normalized to 100. The dependent variable is market leverage, defined as total debt scaled by the market value of assets (TDM). The description of selected variables is given in Internet Appendix Table A1.

We begin by presenting the predictive performance of these three alternative metrics of leverage. Internet Appendix Fig. A7 plots predictability (R^2_{OS}) across our out-of-sample period (2001–2018) for each machine learning model as well as the linear model. The panels in this figure highlight several salient observations. RF has higher R^2_{OS} in all three panels. Focusing on book leverage in Panel B, for most years, RF averages around 20% greater than the LM and around 4% more than GBM. GBM and NNET for all years outperform the LM by around 15%. LASSO and the LM show similar levels of underperformance. Finally, there is a modest downward trend in predictability for all models after 2012; e.g., RF predictability decreases about 12% for TDA between 2012 and 2018. However, despite

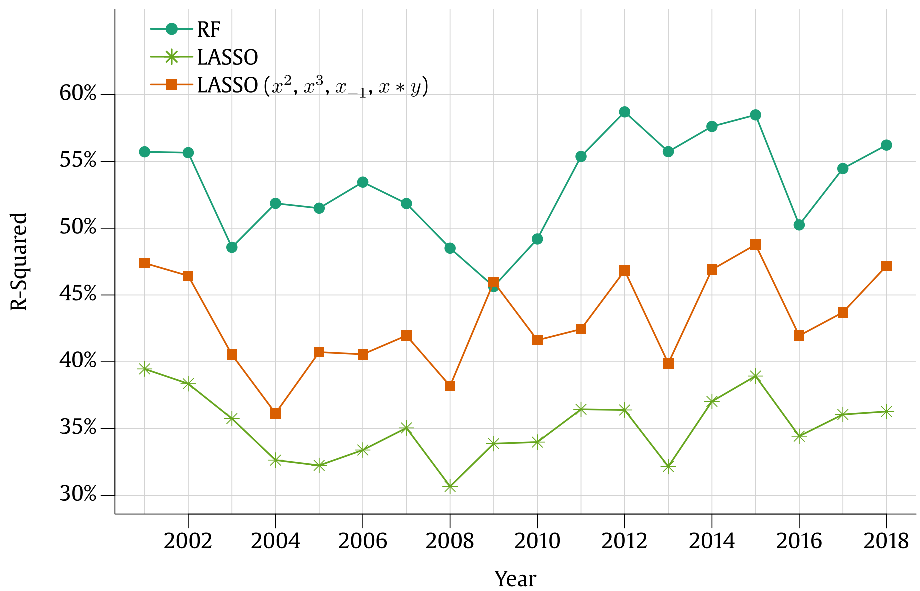

Fig. 5. Out-of-sample performance using augmented, nonlinear LASSO models.
This figure plots the out-of-sample R-squared, R^2_{OS}, for random forests (RFs), least absolute shrinkage and selection operator (LASSO) models, and an augmented, nonlinear LASSO model including squared plus cubed plus lagged plus interaction terms (LASSO (x^2, x^3, x_{-1}, $x * y$)). The out-of-sample period is 2001–2018, and the dependent variable is market leverage, defined as total debt scaled by the market value of assets (TDM).

the decline in predictive performance, RF remains the best performing model over this period.

To further assess the performance of the models, we report the cumulative percentage difference between the MSFE of the linear model and our machine learning models in Internet Appendix Fig. A8. The three plots tell a remarkably consistent story. Similar to Internet Appendix Fig. A5, we observe steady positive slopes for RF, GBM, NNET, and GAM, implying that these models generate forecasts with lower out-of-sample errors every year compared to the LM irrespective of how corporate leverage is quantified. LASSO's relatively flat line in all three panels indicates that this linear method leads to approximately the same forecast errors as LM. LASSO's poor performance highlights the importance of allowing for nonlinearities and interactions, as this method imposes linear parameters after it chooses its variables, and its modeling does not explicitly allow for interactions between variables.

Internet Appendix Fig. A9 displays the variable importance for LDM, TDA, and LDA; we also include TDM for ease of comparison. To save space, we only report the importance of RF. The relevant comparison to TDM is TDA, where the total debt is scaled by the book value of assets, not the market value of assets. Cash emerges as the most important determinant of book leverage, followed by *Z*-Score. These factors represent the third and fourth most important determinant of market leverage (TDM). Overall, the top 5 and 10 drivers of TDA are the same as those for TDM. Inspection of LDM and LDA additionally highlights the similarity of drivers of variables. Industry median leverage, market-to-book, cash, and Z-Score are the main factors driving these variables.

5.4. More on nonlinearities

Our findings suggest that machine learning models that exploit hidden, nonlinear relations between variables are powerful tools in predicting corporate leverage. It is not surprising that a more flexible model fits the data better. A natural question is how improved model fits from machine learning models affect the drivers of firms' leverage dynamics. For which variables does adding nonlinear terms matter? What degree of nonlinearity contributes to predictability? Is, for example, adding squared, cubed, or interactive covariates to a standard linear model sufficient to achieve improvements in predictability comparable to those obtained from machine learning models? What is the theoretical justification for including nonlinear terms in leverage models? In this section, we address these questions.

We begin by assessing the extent to which machine learning models emulate a trivial functional form used in a general linear model. We augment the pool of independent variables in our LASSO model with lagged, quadratic, cubed, and interaction terms. In particular, we estimate the following specifications: (1) squared terms (LASSO (x^2)), (2) squared plus cubed terms (LASSO (x^2, x^3)), (3) interaction terms (LASSO ($x * y$)), (4) squared plus interaction terms (LASSO (x^2, $x * y$)), (5) squared plus cubed plus interaction terms (LASSO (x^2, x^3, $x * y$)), (6) lagged terms (LASSO (x_{-1})), (7) squared plus lagged terms (LASSO (x^2, x_{-1})), (8) squared plus cubed plus lagged terms (LASSO (x^2, x^3, x_{-1})), (9) lagged plus interaction terms (LASSO (x_{-1}, $x * y$)), (10) squared plus lagged plus interaction terms (LASSO (x^2, x_{-1}, $x * y$)), and (11) squared plus cubed plus lagged plus interaction terms (LASSO (x^2, x^3, x_{-1}, $x * y$)), and calculate the corresponding R^2_{OS} values.

Using TDM, Fig. 5 illustrates the R^2_{OS} for our most comprehensive LASSO model (specification (11), described in the prior paragraph) and compares it with our base RF model, and first-order LASSO with no additional terms. Internet Appendix Table A4 reports the R^2_{OS} for specifications (1)–(10) plus our baseline models, RF, and LASSO. The results clearly demonstrate that the machine learning model, RF, continues to significantly outperform augmented LASSO models. That is, machine learning models detect nonlinearities that are far beyond common nonlinear forms used in traditional models. Additionally, Fig. 5 shows that nonlinear LASSO models outperform the first-order LASSO in all of the years in the testing period, if one insists on using a LASSO or OLS model as opposed to ML models. We draw similar conclusions using book leverage (TDA).

These findings potentially explain our earlier conclusions (Section 5.2) that ML models identify a larger set of reliable leverage determinants compared to other studies, such as Frank and Goyal (2009) and Lemmon et al. (2008). Our results suggest that machine learning algorithms are capable of exploiting hidden, unknown nonlinearities between leverage factors that are beyond the simple, common nonlinear forms. As a result, ML models identify a broader set of covariates that influence corporate leverage.

Finally, we ask what the theoretical justification is for including nonlinear terms in leverage models? We argue that the relative importance of capital structure determinants discussed in Section 5.2 behaves differently under diverse economic environments. Prior studies (e.g., Graham and Leary, 2011) show that nonlinearities in the form of interaction terms are important in explaining leverage. To help us better understand the drivers of the interactive effects that machine learning models do not directly reveal, we, in untabulated results, interact our proxies for stock market conditions, debt market conditions, and macroeconomic conditions with the remaining leverage determinants. We then reestimate the R^2_{OS} statistics as well as variable importance for the RF and LASSO models. The findings[5] show that the R^2_{OS} of the machine learning model, RF, as expected, is significantly higher than the R^2_{OS} of the linear model, LASSO. However, since machine learning models by design incorporate hidden interactions implicitly, feeding the interaction terms directly into the models does not improve the R^2_{OS} relative to our baseline specifications without explicit interaction terms. The findings show that the top three important variables (market-to-book, industry leverage and cash for TDM, and cash, Z-Score, and industry leverage for TDA) remain the top three important variables even after adding the macro-level variables directly into the models.

The results also suggest that while the interaction terms are indeed important in explaining capital structure, not all interactions are equally important. Inflation and macro-level growth interactions with industry leverage, cash, and market-to-book seem to be the most relevant ones. Overall, the findings demonstrate the significance of both direct effects of the leverage determinants (individual terms)

[5] These empirical analyses are available upon request.

as well as their indirect effects through the economic environment (interaction terms). That is, the importance of some determinants varies depending on the economic environment. Having shown that leverage and its determinants are linked in complex, nonlinear forms, the next section examines whether this complex structure affects the estimates of how fast firms adjust their leverage to a desired target.

6. Machine learning and leverage adjustment

6.1. Estimating leverage adjustment speed

Machine learning has great applied value when inferring the speed of adjustment to target leverage. Typically, existing empirical capital structure studies treat the forecasting of target leverage as an estimation step, while it is essentially a prediction task and the coefficients in the first stage are merely a means to estimating the fitted values. In a partial adjustment framework, depending on the costs and benefits of rebalancing their capital structure, firms assess how quickly to close any gap between their actual and their target capital structure. The gap is defined as $GAP_{i,t} = E(y_{i,t+1}) - y_{i,t}$, and the partial adjustment model is estimated as follows:

$$\Delta y_{i,t+1} = \lambda GAP_{i,t} + \varepsilon_{i,t+1}. \quad (10)$$

The adjustment speed, λ, allows firms to move only partially towards their target leverage during year t. If managers have target debt ratios and expend efforts to reach them, λ should be strictly greater than zero. In other words, firms should adjust when there is a wedge between the target and the actual leverage ratio. In the presence of market frictions, the adjustment is not instantaneous, and λ is less than one.

We now apply our partial adjustment framework to evaluate a firm's adjustment speed to its target leverage. This step involves the estimation of Eq. (10) as a pooled OLS regression with bootstrapped standard errors to account for the generated regressor (Pagan, 1984). We are particularly interested in assessing whether the superior performance of machine learning in target leverage prediction translates into alternative inferences regarding the rate of leverage adjustment.

Panels A and B of Table 3 report the adjustment speed estimates for both machine learning and linear methods, using TDM and TDA, respectively. We calculate capital structure adjustment speed under different assumptions about the set of target leverage predictors. Columns (1)–(3) use firm-level, industry-level, and macroeconomic-level predictors of target leverage. Columns (4)–(6) add firm fixed effects to these determinants. In our models, fixed effects are controlled by demeaning the variables in the panel data sample.

Table 3
The speed of leverage adjustment.

	RF	LM	LASSO	RF	LM	LASSO
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: TDM						
GAP	0.215***	0.171***	0.161***	0.473***	0.430***	0.428***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	38,069	38,069	38,069	38,069	38,069	38,069
Adjusted R^2	0.062	0.060	0.056	0.188	0.185	0.183
Half-life in years	2.86	3.70	3.94	1.08	1.23	1.24
Panel B: TDA						
GAP	0.150***	0.108***	0.107***	0.345***	0.322***	0.321***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	38,069	38,069	38,069	38,069	38,069	38,069
Adjusted R^2	0.048	0.038	0.037	0.137	0.133	0.133
Half-life in years	4.26	6.07	6.12	1.64	1.78	1.79

This table reports the speed of adjustment estimate of an average firm towards its desired leverage by evaluating the following model:

$$\Delta y_{i,t+1} = \lambda GAP_{i,t} + \varepsilon_{i,t+1},$$

where $GAP_{i,t} = E(y_{i,t+1}) - y_{i,t}$ represents the gap between firm i's actual and its target capital structure. Target leverage ($Ey_{i,\ t+1}$) is predicted using either the best performing machine learning method (RF) or one of the two linear data models (LM and LASSO). In Panel A, the dependent variable is market leverage, defined as total debt scaled by the market value of assets (TDM), and in Panel B, the dependent variable is book leverage, defined as total debt scaled by the book value of assets (TDA). Columns (1)–(3) use firm, industry, and macroeconomic predictors of target leverage. Columns (4)–(6) add firm fixed effects to the set of control variables. *P*-values, which are based on bootstrapped standard errors to account for the generated regressor bias, are reported in parentheses (Pagan, 1984). Symbols ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

We note several important findings. First, machine learning is associated with 8%–40% higher speed of adjustment estimates relative to linear models. Using firm, industry, and macroeconomic determinants as predictors of target market leverage, the speed of adjustment for a typical firm is 17.08% for OLS and 16.14% for LASSO (Columns 2 and 3, Panel A) compared to the corresponding adjustment speed estimate of 21.49% for RF (Column 1, Panel A). Similar conclusions are obtained for book leverage. The speed of adjustment of an average firm ranges from 10.71% to 10.79% using linear models (Columns 2 and 3, Panel B) as opposed to 15.02% using machine learning (Column 1, Panel B). In relative terms, machine learning boosts the speed of adjustment by 33% (40%) relative to LASSO for market (book) leverage. In economic terms, firms converge to their target leverage at a relatively slower speed, with a half-life of about 3.70 and 3.94 (6.07 and 6.12) years for market (book) leverage using OLS and LASSO methods. However, firms adjust at a much faster speed, with a half-life of 2.86 (4.26) years using machine learning. The machine learning estimates suggest that the adjustment speed is relatively faster than most studies document using linear methods.

Second, firm fixed effects play a critical role in the adjustment speed estimation, more than doubling the initial estimates. For example, machine learning yields a speed of adjustment of 21.49% and 15.02% (Column 1, Panels A and B) with no firm fixed effects, using market and book leverage, respectively. The estimates jump to 47.32% and 34.53% with firm fixed effects (Column 4, Panels A and B). The corresponding differences are even more striking using linear methods. The adjustment speed estimates dramatically increase from 16.14% and 17.08% (Columns 3 and 2, Panel A) to 42.75% and 43.04% (Columns 6 and 5, Panel A) for market leverage and from 10.71% and 10.79% (Columns 3 and 2, Panel B) to 32.12% and 32.21% (Columns 6 and 5, Panel B) for book leverage. These findings show the importance of accounting for unobservable firm characteristics in estimating the adjustment speed using either linear or machine learning models.

Third, the adjustment speed estimates using machine learning and incorporating firm fixed effects (Column 4 in Panels A and B) imply that, on average, firms converge to their target relatively fast, closing annually about one third to one half of the leverage gap for market and book leverage, respectively. The estimated half-life in closing the leverage gap using machine-learning-generated targets is only a year for market leverage and 1.6 years for book leverage. This means that firms indeed rebalance away from the unwanted effects of leverage shocks, and the rebalancing occurs at a relatively quicker speed than most studies document.

Fourth, controlling for firm fixed effects, linear models continue to underperform machine learning, although the differences across the models shrink considerably. Machine learning increases the speed of adjustment by 10% (8%) relative to LASSO for market (book) leverage. Overall, our results show that accounting for cross-sectional heterogeneity in leverage dynamics helps partially overcome the inability of linear models to account for complex and nonlinear relations between leverage and its determinants, but it does not fully resolve the issue.

Our results also help reconcile the different estimates of the speed of adjustment documented in the literature by showing that leverage adjustment rates are sensitive to the econometric procedure and the set of predictors employed in extracting the leverage targets. Many studies may have attributed the slower speed of adjustment to elevated adjustment costs prohibiting leverage adjustments when the true target leverage was likely measured with error. Linear data methods overlook nonlinearities and complexities inherent in leverage dynamics creating a considerable downward bias in the resulting estimates.

Our evidence is consistent with Flannery and Hankins (2013), who apply the panel GMM of Blundell and Bond (1998) to improve target leverage estimation. By simultaneously estimating leverage and adjustment speed dynamics and incorporating firm fixed effects, these studies are also able to partially account for complexities in leverage dynamics. Consequently, these methods yield speed of adjustment estimates between 20% and 30%.[6] However, machine learning further boosts adjustment estimates to 35%–50% annually.

6.2. Determinants of leverage adjustment speed

Empirical assessments of leverage adjustment determinants lend support to different factors such as transaction costs (Fischer et al., 1989; Leary and Roberts, 2005; Strebulaev, 2007), macroeconomic conditions (Korajczyk and Levy, 2003; Cook and Tang, 2010), investment (DeAngelo et al., 2011) and cash flow considerations (Faulkender et al., 2012), equity mispricing (Warr et al., 2012), and economic uncertainty (Çolak et al., 2018). To our knowledge, no study systematically tests the entire set of factors that determine the speed with which firms converge to their target leverage. To fill this gap, we use our RF model to analyze and rank the impact of a large set of firm, industry, and macro-level factors on the speed of adjustment. To incorporate the potential factors into the determination of the adjustment speed, we modify Eq. (10) and estimate the following model:

$$\Delta y_{i,t+1} = GAP_{it} + GAP_{it} \times X_{1it} + GAP_{it} \times X_{2it} + \cdots + GAP_{it} \times X_{nit} + \varepsilon_{i,t+1}, \tag{11}$$

where X_1, X_2, $\cdots$, X_n are potential characteristics affecting the adjustment speed. To examine the importance of each factor in explaining the variation in the adjustment speed, we use the 2001–2007 period as the training period and 2008–2018 period as the testing period. For robustness, we repeat our analyses by extending the training period to include the financial crisis period and the results remain qualitatively unchanged. We report the importance of factors for the last year of our sample, 2018.[7] Fig. 6 displays the results. Panels A and B present variable importance results for TDM when the 2001–2007 and 2001–2010 periods are used to train the

[6] Faulkender et al. (2012) and Flannery and Hankins (2013) use a dynamic GMM framework to avoid the Nickell bias (Nickell, 1981) associated with a lagged dependent variable and fixed effects. We avoid the Nickell bias by using only fixed effects, not a lagged dependent variable, as the resolution of this bias has not been addressed in a machine learning framework.

[7] The findings are generally consistent across the other years of the out-of-sample period.

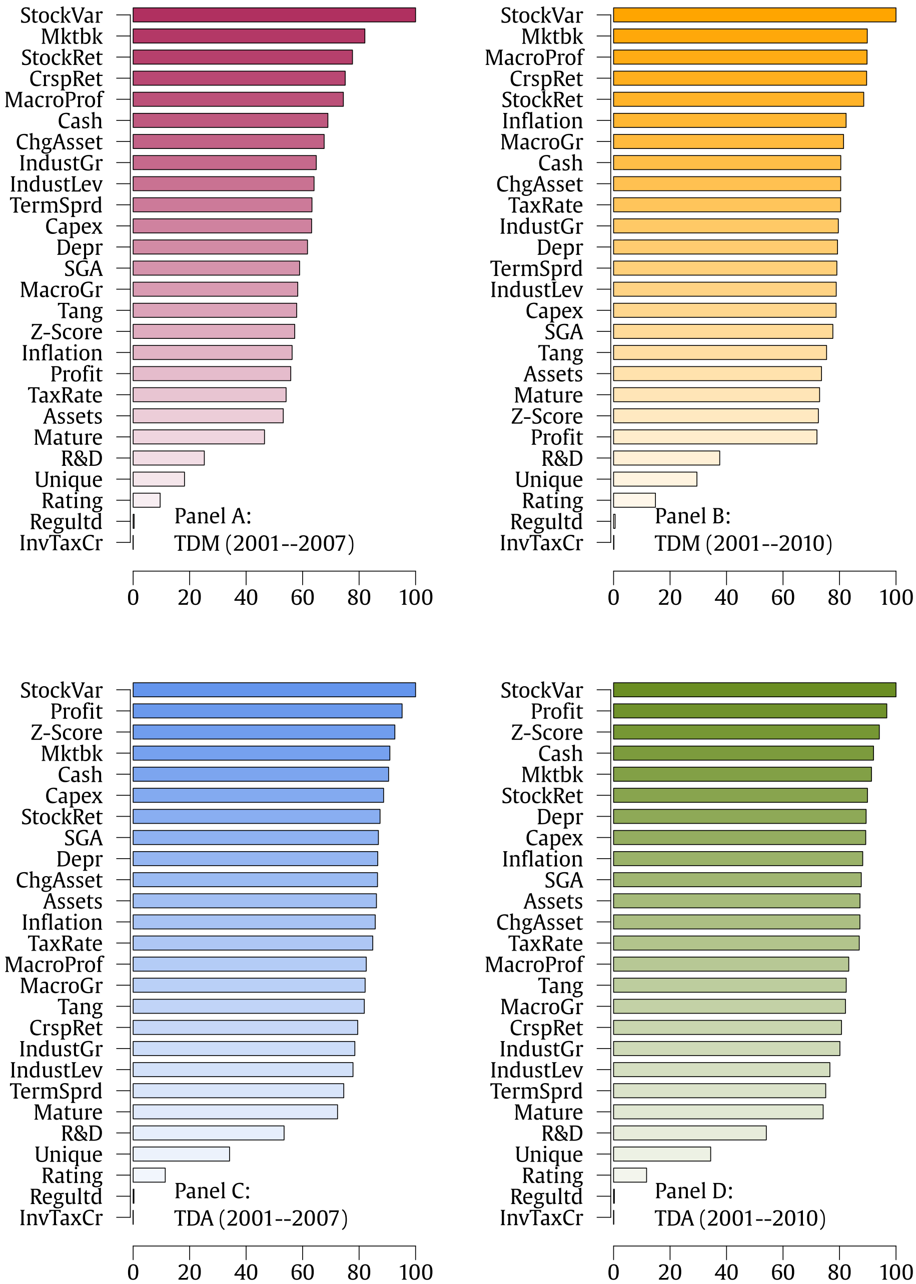

Fig. 6. Final variable importance for adjustment speed.
This figure plots the importance of explanatory variables for estimating the speed of adjustment using a random forest (RF) model. The highest importance value is normalized to 100. The estimates are obtained from training the model over the period 2001–2007 in Panels A and C, and 2001–2010 in Panels B and D. The variable importance is reported for the last year of the sample (2018). The dependent variable is $\text{TDM}_{it} - \text{TDM}_{it-1}$, or $\text{TDA}_{it} - \text{TDA}_{it-1}$ (for variable descriptions, see Internet Appendix Table A1).

RF model, respectively. Panels C and D report the corresponding results for TDA.

Several variables are relevant in determining the speed of adjustment; stock variance emerges as the most prominent factor. This result complements the findings of Altinkiliç and Hansen (2015) and Çolak et al. (2018), who demonstrate that leverage adjustments depend largely on economy-wide uncertainty. We show for the first time that uncertainty at the firm level is also vital for leverage adjustments. Cash considerations are also major drivers of the adjustment speed, consistent with the idea that funding investment or liquidity via a low-cost means of issuance is critical for leverage adjustments (Faulkender et al., 2012). Market-to-book and stock returns are additional relevant factors, supporting the relevance of equity pricing and market timing in affecting leverage adjustments (e.g., Warr et al., 2012; Faulkender et al., 2012). The significance of business cycle variables, such as macro-level profits and growth,

Table 4
Portfolio analysis for differences in firm actions.

	Expected Sign	RF	LM	LASSO
	(1)	(2)	(3)	(4)
Panel A: Market Leverage				
Debt Retirement	(+)	0.062***	0.054***	0.064***
Equity Issue	(+)	0.012***	−0.014***	−0.024***
Debt Issue	(−)	−0.006*	0.040***	0.024***
Equity Retirement	(−)	−0.011**	0.020***	0.047***
Net Debt Issue	(−)	−0.020***	0.094	0.044
Net Equity Issue	(+)	0.008***	−0.002	−0.015***
Leverage Increase	(−)	−0.011**	0.053***	0.018***
Panel B: Book Leverage				
Debt Retirement	(+)	0.058***	0.060***	0.061***
Equity Issue	(+)	0.010**	0.002	0.003
Debt Issue	(−)	−0.026***	−0.010**	−0.021***
Equity Retirement	(−)	−0.012***	−0.009**	−0.008*
Net Debt Issue	(−)	−0.026***	−0.002***	−0.023***
Net Equity Issue	(+)	0.004**	0.010***	0.009***
Leverage Increase	(−)	−0.022***	0.008*	−0.013***

The table presents differences in firm actions between firms with leverage above the target and firms with leverage below the target. Leverage targets are obtained using a random forest (RF) in column (2) and LM and LASSO in columns (3) and (4), respectively. The leverage gap for firm *i* in year *t* is defined as

$$Gap_{it} = L_{it-1} - L^*_{it},$$

where L_{it-1} is realized leverage for firm *i* at the end of year $t-1$, and L^*_{it} denotes the target leverage estimated by the model of choice. Firms are then divided into groups of above target leverage (overleveraged, $Gap_{it} > 0$) and below target leverage (underleveraged, $Gap_{it} < 0$), and average firm actions for each group are calculated. Differences in firm actions between overleveraged and underleveraged firms are then reported in this table. In a given year, debt issue, debt retirement, equity issue, and equity retirement actions refer to the case in which the change in equity or debt, normalized by the book value of assets at the end of the previous period, is greater than 5% of total assets. Net debt issue is debt issue minus debt retirement. Net equity issue is equity issue minus equity retirement. Market (book) leverage increase is a dummy variable that equals one if the change in market (book) leverage is positive, and zero otherwise. Panels A and B use market and book leverage, respectively. Column (1) shows the expected signs of differences in firm actions. Symbols ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

confirms the tendency of firms to rebalance their capital structure when the macro-level conditions are more favorable (e.g., Korajczyk and Levy, 2003; Cook and Tang, 2010). Finally, capital expenditure is similarly a highly ranked determinant, consistent with the relevance of investment considerations for leverage rebalancing decisions (DeAngelo et al., 2011; Dudley, 2012; Elsas et al., 2013).

6.3. *Financing actions*

A firm's issuance and retirement of its securities are observable. This allows us to examine whether the superior predictive performance of machine learning models has implications for corporate actions in the debt and stock markets. Frank and Shen (2019) test a series of linear leverage models and find that these models do not correctly predict actual debt issues and stock sales. How does the prediction of the actual financing actions by machine learning models compare to that of linear models? To answer this question, we use the predicted leverage targets from both machine learning and linear models to identify firms with suboptimal capital structure and track their financing decisions in the subsequent years. According to theory, a firm not at its target leverage should engage in financing actions to adjust towards the target and gradually close its leverage deviation. Since machine learning models, on average, yield more accurate predictions of leverage targets relative to linear models, we expect machine learning predictions of financing actions to be more consistent with the predictions of theory.

We compare the performance of RF, LM, and LASSO models in predicting the expected direction of financing actions for over and underleveraged firms. We empirically test whether capital market transactions in the form of debt retirement, equity issues, debt issues, and equity retirement are consistent with the prediction of the theory (see Frank and Shen, 2019). Firms are initially divided into groups of above-target leverage and below-target leverage. We then calculate average firm actions for each group and test the differences in means. For each financing action and with each target leverage measure, we evaluate whether overleveraged (underleveraged) firms reduce (increase) their leverage more compared to underleveraged (overleveraged) firms. According to theory, overleveraged (underleveraged) firms should rely more on equity (debt) issues and/or debt (equity) retirements to reduce (increase) their leverage and converge to their target. Table 4 presents the results of the differences in means tests. Column (1) provides the expected sign on each financing action under the assumption that each model correctly predicts the leverage targets. Columns (2), (3), and (4) show the results for the RF, LM, and LASSO models, respectively.

The results are strikingly different across methods. Using market leverage in Panel A, all three models correctly predict the cross-sectional patterns of debt retirements. However, linear models make incorrect predictions about the remaining financing actions. The wrong signs from both LM and LASSO are highly statistically significant for equity issues, debt issues, equity retirements, and leverage increases. LASSO also fails to predict net equity issues, with an opposite sign and high statistical significance. In Panel B, where we use book leverage, the performance of linear models improves slightly. Yet, both LM and LASSO still fail to significantly predict the cross-sectional differences in equity issues. Additionally, LM is unable to predict the cross-sectional patterns of leverage increases, which loads with the wrong sign and high statistical significance. In contrast, RF correctly predicts the cross-sectional patterns of each financing action consistently along with high statistical significance in both panels. Overall, these results suggest that machine learning has a greater ability to predict observed financing actions compared to linear models. More noisy estimates of leverage targets by linear models lead to the potentially misleading interpretation that firms do not undertake corrective actions to adjust their leverage. Machine learning, on the other hand, correctly identifies firms with suboptimal capital structure and provides sensible predictions of the cross-sectional differences in firms' financing actions. These findings indicate that targets generated by machine learning conform to theory.

In a series of untabulated analyses, we also evaluate the value implications of suboptimal leverage. To assess whether the market perceives leverage deviations as value-destroying, we evaluate the valuation of firms with suboptimal leverage ratios relative to those that are in the optimal range. We categorize firms as over or underlevered based on leverage targets predicted by machine learning. We then use this indicator variable as our main variable of interest in a regression with Tobin's Q as the dependent variable, while controlling for several relevant firm attributes, such as profitability, size, capital expenditures, R&D spending, sales growth, and age along with year, firm, and industry fixed effects. We find that firms with suboptimal leverage have lower firm value. This finding highlights the possibility that managers attempt to converge to their targets to avoid potentially value-destroying effects of suboptimal capital structure.[8]

7. Robustness tests

Thus far, our study demonstrates that machine learning models that accommodate for nonlinear and interaction effects outperform linear models in predicting corporate leverage. The predictive performance of capital structure theories, however, is also shown to depend on firm conditions and environment (Myers, 2003). Therefore, we test whether machine learning models outperform linear models in predicting target leverage in different subsamples to confirm that results are robust to firm size, time period, and economic environment.

To evaluate the possibility that a particular type of firm may be driving our conclusions, we classify firms every year into (1) small, medium, and large firms based on their size; (2) low-, medium-, and high-growth firms based on their market-to-book; or (3) high-tech and non-high-tech firms based on their industry. We then report R^2_{OS} statistics every year for RF and LM.

Panels A and B in Internet Appendix Fig. A10 and Panel A in Internet Appendix Fig. A11 present R^2_{OS} statistics for size, growth, and technology portfolios, respectively. Results show that the superior predictive performance of machine learning methods carries over into our examined subsamples. RF consistently generates higher R^2_{OS} statistics relative to the linear model for each classification every year. Illustrated in Panel A of Internet Appendix Fig. A10, the difference between the R^2_{OS} of the RF and the LM model ranges from 8% in 2003 to 26% in 2004 for large-cap firms. The difference for small-cap firms ranges from 7% in 2003 to 24% in 2013. Similarly, Panel B of Internet Appendix Fig. A10 and Panel A of Internet Appendix Fig. A11 document significant and large improvements in R^2_{OS} for growth and value firms as well as high-tech and non-high-tech firms. Overall, these results provide convincing evidence that improvement gains in predicting corporate leverage by using machine learning methods are not exclusively an average firm effect. Instead, machine learning methods can be reliably applied to small firms, high-growth firms, or high-tech firms as well as their counterparts.

The trade-off theory predicts that a firm's goal is to maintain an optimal range of debt ratio where the tax benefits and bankruptcy costs of raising debt are balanced. However, empirical studies (Fama and French, 2002; Leary and Roberts, 2005; Flannery and Rangan, 2006, among others) document that firms rebalance their leverage infrequently and not immediately when their debt ratio deviates from the desired level. Dynamic trade-off theories (e.g., Leary and Roberts, 2005) reconcile this observation and argue that the cost of issuing debt prevents firms from continuously rebalancing their capital structure. Firms readjust their debt ratio when the benefits of recapitalization outweigh the issuance costs. As a result, significant capital restructuring in practice occurs infrequently, but debt changes are relatively large when they occur. Consequently, it is meaningful to predict debt ratios during capital refinancing periods because firms are moving to their target at those times.

We examine how accurately machine learning models predict the leverage of firms that implement a major capital adjustment. We follow Danis et al. (2014) to identify years in which firms adjust their capital structure by a large amount. More specifically, we flag a firm in our sample as refinancing if its net long-term debt issuance relative to assets exceeds 3%, and its net payouts (cash dividends plus net share repurchases) relative to total assets exceed the 3% threshold. We then focus on the out-of-sample fit for the two subsamples (refinancing and non-refinancing firms) using our RF and LM estimators. The descriptive statistics for net debt issuance (NetDebt) and net payouts measures (NetPay) are in Panel D of Table 1.

Panel B of Internet Appendix Fig. A11 reports R^2_{OS} statistics. Results show that RF consistently dominates the LM estimator in predicting the debt ratio of both refinancing and non-refinancing firms. RF predictability measured by R^2_{OS} peaks at over 73% for

[8] These results should be interpreted with caution since it is not clear whether the firm value drops due to managers not making proactive corrections to their capital structure, or because they are subject to more frequent and larger shocks to their capital structure.

refinancing firms in 2015. Further, R^2_{OS} gains relative to LM range from 9% in 2009 to 36% in 2012 for firms that are adjusting their capital structure. In summary, over the entire sample, RF's average R^2_{OS} for both refinancing and non-refinancing firms are relatively similar, approximately 55% and 53%, respectively. The takeaway from these results is that machine learning models such as RF possess greater predictive power than LM for firms when they undergo major recapitalizations of their capital structure and move closer to their target leverage ratio.

8. Conclusion

In this paper, we undertake the first comparative analysis between machine learning and linear models to identify key determinants of capital structure dynamics, estimate target leverage and adjustment speed, and predict financing actions. Our primary contributions are threefold. First, relative to linear models, machine learning selects a broader set of leverage determinants. Second, machine learning algorithms provide more accurate target measurement leading to a faster speed of adjustment and a more theoretically-consistent prediction of observed firms' financing patterns. Third, machine learning models allow us to rank the relative importance of factors affecting the speed of adjustment towards target leverage. Overall, the superior performance of machine learning models relative to linear counterparts reinforces the importance of incorporating interactions and nonlinearities in modeling capital structure dynamics.

Firms only sporadically adjust leverage, payout, and investment due to various frictions such as transaction costs and irreversibility, introducing complex nonlinearities in the data. When this occurs, prediction errors can be costly, resulting in erroneous conclusions about firm behavior. In the context of financing choices, we show that machine learning models are well-suited to predict lumpy firm dynamics and offer robust performance gains compared to linear models. Machine learning models, therefore, should be in the toolbox of scholars and practitioners examining corporate decisions involving intermittent adjustments.

Finally, the machine learning methodologies that we highlight in this paper (RF, GBM, and NNET) all perform admirably when faced with potentially unknown nonlinear and interaction effects. It is important to note that these effects do not need to be identified a priori as would be the case with traditional linear models and/or a LASSO. We stress that, although the RF models tend to outperform the GBM and NNET in modeling leverage, there are no optimality criteria that suggest that RF is always consistently superior to GBM and NNET. Without question, there are (perhaps unknown) scenarios that would show NNET (particularly deep learning networks) and/or GBM models to be superior to the random forest. Therefore, we introduce all models in this paper and advocate for scholars and practitioners to try a battery of machine learning methods in any given situation.

Appendix A. Supplementary tables and figures

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jcorpfin.2021.102073.

References

Altinkiliç, O., Hansen, R.S., 2015. Are there economies of scale in underwriting fees? Evidence of rising external financing costs. Rev. Financ. Stud. 13 (1), 191–218.
Altman, E.I., 1968. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. J. Financ. 23 (4), 589–609.
Altman, N., Krzywinski, M., 2017. Ensemble methods: bagging and random forests. Nat. Methods 14 (10), 933–934.
Archer, K.J., Kimes, R.V., 2008. Empirical characterization of random forest variable importance measures. Comput. Statist. Data Anal. 52 (5), 2249–2260.
Baker, M., Wurgler, J., 2002. Market timing and capital structure. J. Financ. 57 (1), 1–32.
Bali, T.G., Goyal, A., Huang, D., Jiang, F., Wen, Q., 2020. The Cross-Sectional Pricing of Corporate Bonds Using Big Data and Machine Learning. Unpublished working paper. Georgetown University.
Barclay, M.J., Smith, C.W., Watts, R.L., 1995. The determinants of corporate leverage and dividend policies. J. Appl. Corpor. Finance 7 (4), 4–19.
Beck, M.W., 2018. NeuralNetTools: visualization and analysis tools for neural networks. J. Stat. Softw. 85 (11), 1–20.
Bianchi, D., Büchner, M., Tamoni, A., 2020. Bond risk premiums with machine learning. Rev. Financ. Stud. 34 (2), 1046–1089.
Blundell, R., Bond, S.R., 1998. Initial conditions and moment restrictions in dynamic panel data models. J. Econ. 87 (1), 115–143.
Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140.
Breiman, L., 2001. Random forests. Mach. Learn. 45 (1), 5–32.
Bubna, A., Das, S.R., Prabhala, N., 2020. Venture capital communities. J. Financ. Quant. Anal. 55 (2), 621–651.
Childs, P.D., Mauer, D.C., Ott, S.H., 2005. Interactions of corporate financing and investment decisions: the effects of agency conflicts. J. Financ. Econ. 76 (3), 667–690.
Çolak, G., Gungoraydinoglu, A., Öztekin, Ö., 2018. Global leverage adjustments, uncertainty, and country institutional strength. J. Financ. Intermed. 35 (Part A), 41–56.
Cook, D.O., Tang, T., 2010. Macroeconomic conditions and capital structure adjustment speed. J. Corp. Finan. 16 (1), 73–87.
Core Team, R., 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Danis, A., Rettl, D.A., Whited, T.M., 2014. Refinancing, profitability, and capital structure. J. Financ. Econ. 114 (3), 424–443.
DeAngelo, H., Roll, R.W., 2015. How stable are corporate capital structures? J. Financ. 70 (1), 373–418.
DeAngelo, H., DeAngelo, L., Whited, T.M., 2011. Capital structure dynamics and transitory debt. J. Financ. Econ. 99 (2), 235–261.
Ductor, L., Fafchamps, M., Goyal, S., van der Leij, M.J., 2014. Social networks and research output. Rev. Econ. Stat. 96 (5), 936–948.
Dudley, E., 2012. Capital structure and large investment projects. J. Corp. Finan. 18 (5), 1168–1192.
Efron, B., Tibshirani, R.J., 1994. An Introduction to the Bootstrap. Monographs on Statistics and Applied Probability. Chapman and Hall/CRC, Boca Raton, FL.
Elsas, R., Florysiak, D., 2015. Dynamic capital structure adjustment and the impact of fractional dependent variables. J. Financ. Quant. Anal. 50 (5), 1105–1133.

Elsas, R., Flannery, M.J., Garfinkel, J.A., 2013. Financing major investments: information about capital structure decisions. Rev. Finance 18 (4), 1341–1386.
Erel, I., Stern, L.H., Tan, C., Weisbach, M.S., 2021. Selecting directors using machine learning. Rev. Financ. Stud. 34 (7), 3226–3264.
Fama, E.F., French, K.R., 2002. Testing trade-off and pecking order predictions about dividends and debt. Rev. Financ. Stud. 15 (1), 1–33.
Faulkender, M., Flannery, M.J., Watson Hankins, K., Smith, J.M., 2012. Cash flows and leverage adjustments. J. Financ. Econ. 103 (3), 632–646.
Fischer, E.O., Heinkel, R.L., Zechner, J., 1989. Dynamic capital structure choice: theory and tests. J. Financ. 44 (1), 19–40.
Flannery, M.J., Hankins, K.W., 2013. Estimating dynamic panel models in corporate finance. J. Corp. Finan. 19, 1–19.
Flannery, M.J., Rangan, K.P., 2006. Partial adjustment toward target capital structures. J. Financ. Econ. 79 (3), 469–506.
Frank, M.Z., Goyal, V.K., 2009. Capital structure decisions: which factors are reliably important? Financ. Manag. 38 (1), 1–37.
Frank, M.Z., Shen, T., 2019. Corporate capital structure actions. J. Bank. Financ. 106, 384–402.
Friedman, J.H., 2001. Greedy function approximation: a gradient boosting machine. Ann. Stat. 29 (5), 1189–1232.
Garson, D.G., 1991. Interpreting neural-network connection weights. Artif. Intell. Expert 6 (4), 46–51.
Goh, A.T.C., 1995. Back-propagation neural networks for modeling complex systems. Artif. Intell. Eng. 9 (3), 143–151.
Goyal, A., Welch, I., 2003. Predicting the equity premium with dividend ratios. Manag. Sci. 49 (5), 639–654.
Graham, J.R., Harvey, C.R., 2001. The theory and practice of corporate finance: evidence from the field. J. Financ. Econ. 60 (2), 187–243.
Graham, J.R., Leary, M.T., 2011. A review of empirical capital structure research and directions for the future. Ann. Rev. Financ. Econ. 3 (1), 309–345.
Gu, S., Kelly, B.T., Xiu, D., 2020. Empirical asset pricing via machine learning. Rev. Financ. Stud. 33 (5), 2223–2273.
Harris, M., Raviv, A., 1991. The theory of capital structure. J. Financ. 46 (1), 297–355.
Hastie, T., Tibshirani, R.J., Friedman, J.H., 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Series in Statistics, Springer, New York, NY.
Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: an investment approach. Rev. Financ. Stud. 28 (3), 650–705.
Hovakimian, A., Hovakimian, G., Tehranian, H., 2004. Determinants of target capital structure: the case of dual debt and equity issues. J. Financ. Econ. 71 (3), 517–540.
Huang, R., Ritter, J.R., 2009. Testing theories of capital structure and estimating the speed of adjustment. J. Financ. Quant. Anal. 44 (2), 237–271.
James, G., Witten, D.M., Hastie, T., Tibshirani, R.J., 2013. An Introduction to Statistical Learning: With Applications in R. Springer Texts in Statistics, Springer, New York, NY.
Jensen, M.C., 1986. Agency costs of free cash flow, corporate finance, and takeovers. Am. Econ. Rev. 76 (2), 323–329.
Kale, J.R., Noe, T.H., Ramìrez, G.G., 1991. The effect of business risk on corporate capital structure: theory and evidence. J. Financ. 46 (5), 1693–1715.
Kayhan, A., Titman, S., 2007. Firms' histories and their capital structures. J. Financ. Econ. 83 (1), 1–32.
Kayo, E.K., Kimura, H., 2011. Hierarchical determinants of capital structure. J. Bank. Financ. 35 (2), 358–371.
Korajczyk, R.A., Levy, A., 2003. Capital structure choice: macroeconomic conditions and financial constraints. J. Financ. Econ. 68 (1), 75–109.
Kuhn, M., Johnson, K., 2013. Applied Predictive Modeling. Springer, New York, NY.
Leary, M.T., Roberts, M.R., 2005. Do firms rebalance their capital structures? J. Financ. 60 (6), 2575–2619.
Lemmon, M.L., Roberts, M.R., Zender, J.F., 2008. Back to the beginning: persistence and the cross-section of corporate capital structure. J. Financ. 63 (4), 1575–1608.
Li, K., Feng, M., Shen, R., Yan, X., 2021. Measuring corporate culture using machine learning. Rev. Financ. Stud. 34 (7), 3265–3315.
Lockhart, G., 2014. Credit lines and leverage adjustments. J. Corp. Finan. 25, 274–288.
Myers, S.C., 1977. Determinants of corporate borrowing. J. Financ. Econ. 5 (2), 147–175.
Myers, S.C., 1984. The capital structure puzzle. J. Financ. 39 (3), 574–592.
Myers, S.C., 2003. Financing of corporations. In: Constantinides, G.M., Harris, M., Stulz, R.M. (Eds.), Corporate Finance, Volume 1A of Handbook of the Economics of Finance, Chapter 4, Pages 215–253. Elsevier.
Myers, S.C., Majluf, N.S., 1984. Corporate financing and investment decisions when firms have information that investors do not have. J. Financ. Econ. 13 (2), 187–221.
Nickell, S.J., 1981. Biases in dynamic models with fixed effects. Econometrica 49 (6), 1417–1426.
Öztekin, Ö., 2015. Capital structure decisions around the world: which factors are reliably important? J. Financ. Quant. Anal. 50 (3), 301–323.
Öztekin, Ö., Flannery, M.J., 2012. Institutional determinants of capital structure adjustment speeds. J. Financ. Econ. 103 (1), 88–112.
Pagan, A., 1984. Econometric issues in the analysis of regressions with generated regressors. Int. Econ. Rev. 25 (1), 221–247.
Rajan, R.G., Zingales, L., 1995. What do we know about capital structure? Some evidence from international data. J. Financ. 50 (5), 1421–1460.
Rauh, J.D., Sufi, A., 2012. Explaining corporate capital structure: product markets, leases, and asset similarity. Rev. Finance 16 (1), 115–155.
Strebulaev, I.A., 2007. Do tests of capital structure theory mean what they say? J. Financ. 62 (4), 1747–1787.
Strobl, C., Boulesteix, A., Kneib, T., Augustin, T., Zeileis, A., 2008. Conditional variable importance for random forests. BMC Bioinform. 9 (1), 307–317.
Tibshirani, R.J., 1996. Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B Methodol. 58 (1), 267–288.
Titman, S., Wessels, R., 1988. The determinants of capital structure choice. J. Financ. 43 (1), 1–19.
Venables, W.N., Ripley, B.D., 1997. Modern Applied Statistics with S-PLUS. Statistics and Computing, Springer, New York, NY.
Warr, R.S., Elliott, W.B., Koëter-Kant, J., Öztekin, Ö., 2012. Equity mispricing and leverage adjustment costs. J. Financ. Quant. Anal. 47 (3), 589–616.
Welch, I., 2004. Capital structure and stock returns. J. Polit. Econ. 112 (1), 106–132.