Enhancing Real Estate Investment Trust Return Forecasts using Machine Learning

Kahshin Leow* and Thies Lindenthal

University of Cambridge

*Corresponding author: Kahshin Leow, Correspondence Address: 17 Mill Lane, Cambridge CB2 1RX, United Kingdom; e-mail: kahshin.leow@kings.cam.ac.uk

Abstract

We extend the emerging literature on machine learning empirical asset pricing by analyzing a comprehensive set of return prediction factors for real estate investment trusts (REITs). We show that machine learning models are superior to traditional ordinary least squares models and find that REIT investors experience significant economic gains when using machine learning forecasts. In particular, we show that REITs are more predictable than stocks and that their higher predictability is stable over time and across industries.

1 Introduction

In this study, we build on the pioneering work of Gu et al. (2020), who combine a broad repertoire of machine learning methods with modern empirical asset pricing research to understand the dynamics of market risk premiums for U.S. stock returns. Their results suggest that machine learning improves the description of expected returns. They find that portfolio performance improves most prominently among the more sophisticated machine learning models due in large part to the non-linear predictor interactions missed by simpler methods. Researchers have replicated Gu et al.'s (2020) work in several specific contexts. For example, Bianchi et al. (2021) employ machine learning methods to predict U.S. Treasury bond returns and find strong statistical evidence in favor of extreme trees and neural networks. Leippold et al. (2022) add to the burgeoning financial machine learning literature by expanding previous findings to China, where they find that the predictability of the Chinese stock market is a fewfold higher than that of the U.S. stock market (Gu et al., 2020). Nonetheless, researchers have yet to apply the latest machine learning techniques to real estate investment trusts (REITs). Given the fundamental differences between stocks and real estate, such research is warranted.

Our study is motivated by two questions. First, do the findings of Gu et al. (2020), Bianchi et al. (2021), and Leippold et al. (2022) apply to the real estate market? If true, the implication is that machine learning will aid in solving practical problems such as market timing, portfolio choice, and risk management, justifying its role in public real estate markets.

Second, owing to the fundamental underlying differences between stocks and real estate, are there differences in the predictability of stock and real estate returns? One hypothesis is that the larger heterogeneity in the stock market makes it a more natural candidate for machine learning techniques than REITs. Another hypothesis cited by Nelling and Gyourko (1998), who find that REITs have lower predictability than stock returns, is that the REIT market has a smaller sample of firms than the general stock market, a numerical difference that has only increased since 1998.

We conduct a large-scale empirical analysis investigating 486 REITs over 1990 to 2021. Our predictor set includes 94 firm-level characteristics for each REIT, the interactions of each characteristic with eight macroeconomic time series variables, and 17 sector dummy variables, totaling 863 baseline signals. Our benchmark ordinary least squares (OLS) regression models using only size and the book-to-market ratio as predictors (OLS-2) and size, the book-to-market ratio, and 12-month momentum as predictors (OLS-3) generate out-of-sample R^2s of 0.36 and 0.31 percent per month, respectively. When we expand the OLS panel model to include our set of over 800 predictors, predictability disappears, as evidenced by the R^2 dropping into negative territory. With so many parameters to estimate, this is not surprising.

Our first evidence that machine learning aids in return prediction in the real estate market emerges from the fact that principal component regression, which reduces the dimension of the predictor set to a few linear combinations of predictors, pulls the out-of-sample R^2 into positive territory at 0.28 percent. The least absolute shrinkage and selection operator (LASSO) and elastic net regularization (ENet) models, which use parameter shrinkage and variable selection, further raise the out-of-sample R^2s to 2.49 and 3.37 percent, respectively.

Next, we expand the model to accommodate non-linear predictive relationships using regression trees and neural networks. By allowing for non-linearities, this further improves the predictions. We find that random forests, gradient-boosted regression trees, and extremely randomized trees yield out-of-sample R^2s of 2.71, 2.70, and 4.52 percent, respectively. The best neural network model produces an out-of-sample R^2 of 5.01 percent. Similar to Gu et al. (2020) and Bianchi et al. (2021), we also find that shallow learning outperforms deeper learning. In the case of REITs, neural network performance peaks at one hidden layer, contrary to stock performance, which peaks at three hidden layers (Gu et al., 2020), but similar to bond performance (Bianchi et al., 2021).

Finally, we compare the predictive performance of the machine learning techniques between real estate and stocks. In Gu et al. (2020), their best neural network (NN3) has a monthly out-of-sample R^2 of 0.40 percent, whereas our NN1's R^2 of 5.01 percent is more than 10 times higher. Gu et al.'s (2020) best regression tree model (gradient-boosted regression trees) has a monthly out-of-sample R^2 of 0.34 percent, a much lower figure than our best regression tree model's R^2 of 4.52 percent (extremely randomized trees). Noting that Gu et al.'s (2020) empirical analysis of the U.S. stock market spans 1957 to 2016, while our study period for REITs is from 1990 to 2021, we re-run the analysis on the stock market using the same time period (1990–2021). The qualitative conclusions remain unchanged: REIT returns are more predictable than stock returns when using machine learning methods.

The R^2s ranging from 3 to 5 percent suggest real-world economic benefits. We compare a long-only value-weighted portfolio of REITs with a long-only portfolio comprising REITs with the highest machine learning forecasts. The former has a Sharpe ratio of 0.49, while the latter has a

Sharpe ratio of 0.60, an outperformance of 22 percent. If we look at portfolios constructed by mean-variance optimization, a long-short sample-based mean-variance portfolio has a Sharpe ratio of 0.18, while a long-short mean-variance portfolio incorporating machine learning forecasts has a Sharpe ratio of 0.54, a staggering outperformance of 200 percent. In short, if one has predictive ability, it can lead to significant economic gains.

2 Literature review

There is a rich history in the finance literature for predicting returns in the cross-section and time series (e.g., Amihud, 2002; Ang et al., 2006; Asness et al., 2000; Bazdresch et al., 2014; Chordia et al., 2001; Palazzo, 2012). Within real estate, a large strand of the literature studies the predictability of REIT returns in the time series, including Liu and Mei (1992), Li and Wang (1995), Nelling and Gyourko (1998), Ling et al. (2000), Lin et al. (2009), and Serrano and Hoesli (2010). Another stream investigates the predictability of REIT returns in the cross-section, including Chui et al. (2003), Ooi et al. (2009), Blau et al. (2011), and Giacomini et al. (2015).

Of all this research, Liu and Mei (1992) are the first to find that REITs are more predictable than stocks and bonds. Using a multifactor latent variable model, they find that an equally weighted equity REIT index is more predictable than a value-weighted stock index, a small cap stock index, and a portfolio of long-term U.S. government bonds. They also find that expected excess returns on the REIT index move more closely with those of small cap stocks and much less with those of government bonds. However, Li and Wang (1995) arrive at the opposite conclusion. They observe that the samples used in previous studies are limited in size or include only equity REITs. By including all the REITs available on CRSP tapes, they ensure their results are free of survivorship bias and find no evidence that REIT returns are more predictable than stock returns, contrary to Liu and Mei (1992). They also show that in a general two-factor asset pricing framework, the REIT market is integrated with the stock market. Using more up-to-date data, Ling et al. (2000) find that REIT returns are far less predictable out-of-sample than in-sample, with the inability to forecast out-of-sample particularly true in the 1990s. Serrano and Hoesli (2010) turn the tables again by re-examining the topic using ARMA and ARMA-EGARCH models and daily returns data. They find that REIT returns are generally more predictable than stock returns, especially in countries with mature and well-established REIT regimes.

While machine learning methods have been used in the real estate literature, they have been mostly limited to the prediction of property values (e.g., Baldominos et al., 2018; Lindenthal & Johnson, 2021; Pai & Wang, 2020; Viriato, 2019). Our work adds to these strands of the literature by applying machine learning techniques to the predictability of REIT returns. To the best of our knowledge, ours is the first comprehensive study of REIT returns using machine learning techniques.

The pioneering work of Gu et al. (2020) at the intersection of finance and machine learning deserves special mention. Before the publication of their paper, machine learning methods appeared only sporadically in the asset pricing literature. The vast majority of those studies apply a single machine learning technique. For example, Rapach et al. (2013) apply LASSO to predict global equity market returns, Moritz and Zimmermann (2016) apply tree-based models to portfolio sorting, and Kelly et al. (2019) use dimension reduction methods to estimate and test factor pricing models. Gu et al. (2020) are the first to apply a wide repertoire of over 10 techniques from various branches of the machine learning literature to predict asset prices. Furthermore, the empirical analysis of Gu et al. (2020) is ambitious in both breadth and depth. Their predictor set includes 94 characteristics for each stock, the interactions of each characteristic with eight aggregate time series variables, and 74 industry dummy variables, giving them over 900 baseline signals altogether. Their time series spans 1957 to 2016. No prior work can match Gu et al. (2020) in terms of predicting stock returns in the cross-section and time series. While we cannot match Gu et al. (2020) in terms of historical depth, as the modern REIT era did not begin until 1990, we can match them in terms of cross-sectional breadth using the same 94 characteristics based on Green et al. (2017) and same eight aggregate time series variables based on Welch and Goyal (2008). We also expand Gu et al.'s (2020) toolkit by adding another machine learning technique called extremely randomized trees, which proves to be highly valuable in predicting REIT returns.

3 Data and methodology

We obtain monthly total returns from the CRSP for all REITs and stocks listed on the NYSE, AMEX, and NASDAQ. Our sample for REITs begins in January 1990 and ends in December 2021, while our sample for stocks begins in January 1963 and ends in December 2021. The number of unique REITs in our sample is 486, while the number of unique stocks in our sample is 32,217. We also obtain the one-month Treasury bill rate to calculate individual excess returns.[1]

In addition, we assemble a large collection of the 94 predictive characteristics documented in Green et al. (2017), of which 61 are updated annually, 13 are updated quarterly, and 20 are updated monthly.[2] We include dummies corresponding to the first two digits of the Standard Industrial Classification (SIC) codes. To avoid outliers, we cross-sectionally rank all the continuous firm-level characteristics period-by-period and map them into the [-1,1] interval following Kelly et al. (2019), Gu et al. (2020), and Leippold et al. (2022). Appendix A provides the details of all the firm-level characteristics.

We also assemble eight macroeconomic predictors following the variable definitions in Welch and Goyal (2008): the dividend-to-price ratio (*dp*), earnings-to-price ratio (*ep*), book-to-market ratio (*bm*), net equity expansion (*ntis*), Treasury bill rate (*tbl*), term spread (*tms*), default spread (*dfy*), and stock variance (*svar*).[3] Appendix B provides the details of all the

---

[1] The monthly one-month Treasury bill rate is available from Kenneth French's website, https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

[2] The data are available from Dacheng Xiu's website, https://dachxiu.chicagobooth.edu/. To avoid forward-looking bias, monthly characteristics are delayed by at most one month, quarterly characteristics have at least a four-month lag, and annual characteristics have at least a six-month lag. Therefore, to predict returns in month $t + 1$, we use the most recent monthly characteristics at the end of month t , most recent quarterly data by the end of $t - 4$, and most recent annual data by the end of $t - 6$. We replace the characteristics missing from Xiu's dataset with the cross-sectional median in each month for each REIT and stock.

[3] The monthly data are available from Amit Goyal's website, https://sites.google.com/view/agoyal145.

macroeconomic variables used in this study.

Throughout our analysis, we adopt a general additive prediction error model to describe the relationship between a REIT's excess return[4] and its corresponding predictors:

$$r_{i,t+1} = E_t\left[r_{i,t+1}\right] + \epsilon_{i,t+1}. \tag{1}$$

We further assume the conditional expectation of the ith REIT's excess return $r_{i,t+1}$ given the information available in period t to be a function of a set of predictors:

$$E_t\left[r_{i,t+1}\right] = g(z_{i,t}), \tag{2}$$

where $z_{i,t}$ is the baseline set of firm-level predictors, REITs are indexed by $i = 1, ..., N_t$, and months are indexed by $t = 1, ..., T$. The functional form of $g(.)$ is left unspecified and it depends on z only through $z_{i,t}$. Thus, our prediction model does not use information from before month t or from individual REITs other than the ith REIT. Our goal is to search for the prediction model from a set of machine learning candidates that provides the best predictive performance. The vector of predictors, $z_{i,t}$, consists of the ith REIT's characteristics, the interaction terms between these characteristics, the macroeconomic predictors, and a set of dummy variables, which can be represented as

$$Z_{i,t} = \begin{pmatrix} c_{i,t} \\ x_t \otimes c_{i,t} \\ d_{i,t} \end{pmatrix}, \tag{3}$$

where $c_{i,t}$ is a 94×1 vector of the REIT-level characteristics, x_t is an 8×1 vector of the macroeconomic predictors, $d_{i,t}$ is a 17×1 vector of the dummy variables, and $\otimes$ denotes the Kronecker product. Hence, the total number of covariates in $z_{i,t}$ is $94 \times (8 + 1) + 17 = 863$.

In total, we consider 11 machine learning methods and three simple linear models. Specifically, we include OLS regression; OLS using only size and the book-to-market ratio as predictors (OLS-2); OLS using only size, the book-to-market ratio, and 12-month momentum as predictors (OLS-3); principle component regression; LASSO; ENet; random forests; gradient-boosted regression trees;

[4] Like Gu et al. (2020) and Leippold et al. (2022), we focus on measuring conditional expected returns in excess of the risk-free rate. Some studies in the finance literature have traditionally referred to this quantity as the "risk premium," as it is the compensation that investors demand for bearing investment risk.

extremely randomized trees; and neural networks with one to five layers (NN1–NN5).

To benchmark the predictive power of the models, we perform a 50–50 training–testing split, dividing our data into two disjoint periods while maintaining the temporal ordering: the training sample (1990–2005) and the testing sample (2006–2021). We use the training sample to estimate the model parameters. The testing sample contains the next 12 months of data after the training sample ends. These data, which never enter the model estimation, are used to test our models' predictive performance. Since machine learning models are computationally intensive, we adopt a sample splitting scheme, as in Gu et al. (2020) and Leippold et al. (2022), by refitting the prediction models annually instead of monthly. When we refit a model, the training sample size is increased by one year, but we maintain the same 12-month size for the testing period, which continues to roll forward to include the next 12 months. We do not require a validation sample, as we do not perform any hyperparameter optimization, following Elkind et al. (2022). Default hyperparameters are used where possible. This forms the lower bound of performance for our machine learning models. Appendices C and D provide further details of the prediction models and their respective hyperparameters. All training is executed with open source libraries on an Apple M1 Ultra chip with a 20-core CPU and a single 48-core GPU.

4 Empirical analysis

We start by exploring our models' predictive performance using out-of-sample predictive R^2.

4.1 Out-of-sample predictability

As in Gu et al. (2020), we rely on the non-demeaned out-of-sample predictive R^2 to allow for a direct comparison with their results for the U.S. stock market. For a given prediction model S, this non-demeaned measure is defined as

$$R^2_{oos,S} = 1 - \frac{\Sigma_{(i,t)\epsilon T}(r_{i,t+1} - \hat{r}^{(S)}_{i,t+1})^2}{\Sigma_{(i,t)\epsilon T} r^2_{i,t+1}}$$

where T denotes the set of predictions assessed on the testing sample. $R^2_{oos,S}$ pools the prediction errors across REITs and over time into a grand panel-level assessment of each prediction model S.

4.2 Full sample analysis

Table 1 compares the out-of-sample predictive R^2 of the machine learning techniques. The first column of Table 1 reports the R^2_{oos} for the entire pooled sample of REITs. The OLS model using all 800+ features produces an R^2_{oos} of -6.89 percent. Simple OLS cannot handle so many predictors, which is unsurprising, as the lack of regularization leaves OLS highly susceptible to overfitting. However, restricting OLS to a sparse parameterization, either by forcing the model to include only two or three covariates (OLS-2 and OLS-3) or by penalizing the specification with LASSO and ENet, generates a substantial improvement over the full OLS model ($R^2_{oos,S}$ of 0.36, 0.31, 2.49, and 3.37 percent, respectively).

Regularizing the linear model using dimension reduction techniques fails to outperform the simpler models such as OLS-2 and OLS-3 despite generating an improvement over the full OLS model, with an R^2_{oos} of 0.28 percent. This contrasts with Gu et al. (2020), whose principal component regression model is one of their best-performing linear models. A possible explanation is that we use a default number of five components for our principal component regression model, while Gu et al.'s (2020) principal component regression models allow hyperparameter tuning that varies the number of principal components from one to 100 for their PCR models.

The regression tree models increase the R^2_{oos} even further. The random forests, gradient-boosted regression trees, and extremely randomized trees models have $R^2_{oos,S}$ of 2.71, 2.70, and 4.52 percent, respectively. Such an improvement demonstrates the superiority of the machine learning methods in capturing the complex interactions between predictors, as emphasized by Gu et al. (2020), Bianchi et al. (2021), and Leippold et al. (2022). Interestingly, our results for the regression trees are more in common with the findings of Bianchi et al. (2021) for the bond market than the findings of Gu et al. (2020) for the stock market. For stocks, the gradient-boosted regression trees are the best-performing regression tree. Conversely, for REITs and bonds, the extremely randomized trees model is the best-performing regression tree, while the gradient-boosted regression trees rank the worst. This is the first, but not the last, indication in our study that REITs exhibit more similarities to bonds than to stocks.

Neural networks are by far the best-performing non-linear method and the best predictor overall. The $R^2_{oos}$is 5.01 percent for NN1. This result highlights the value of incorporating complex predictor interactions, which are embedded in neural network models but missed by OLS and other machine learning techniques. These results also show that the benefits of "deep" learning are limited for REITs, concurring with the findings for stocks and bonds. This is likely an artifact of the small amount of data for REITs and tiny signal-to-noise ratio for our return prediction problem compared with the types of non-financial settings in which deep neural networks thrive thanks to large datasets and strong signals such as visual recognition.[5] Interestingly, our neural network findings concur more closely with Bianchi et al. (2021), who find that predictive performance for bonds peaks at one hidden layer, whereas Gu et al. (2020) find that predictive performance for stocks peaks at three hidden layers. Once again, we see hints that REITs behave more closely to bonds than to stocks.

4.3 Predictability over time

In this section, we explore the time variations in the out-of-sample $R^2_{oos}$of our models. Table 2 shows the average monthly R^2_{oos} for all the models by calendar year. The blue bars in Figure 1 show the average monthly R^2_{oos} for OLS-2 and our top three machine learning models (ENet, extremely randomized trees and NN1), sorted by calendar year. We can see a

---

[5] In 2012, AlexNet, a neural network model, made history when it won the prestigious ImageNet Large Scale Visual Recognition Challenge with an error rate of 15.3 percent, more than 10.8 percentage points lower than the runner-up. It contains eight neural network layers. In 2015, Microsoft Research Asia outperformed AlexNet and won the ImageNet contest using over 100 layers in its convolutional neural network architecture.

significant drop in the R_{oos}^2 in 2007 and 2008, with the exception of NN1. We conjecture that the cause of this drop lies in the U.S. subprime mortgage crisis that began in 2007. This points out a possible weakness of machine learning techniques. Their predictive ability can be vulnerable to unexpected systematic risk, such as, in this case, the subprime crisis. Conversely, machine learning models can have fantastic breakout years, such as 2020 in which they exhibit double-digit positive R_{oos}^2 despite it being a terrible year for the real estate market. What happens when we remove these fantastic breakout years from the test sample? The last column of Table 2 shows the overall predictive R_{oos}^2 after excluding 2020 and 2021. The results are qualitatively unchanged; the machine learning models, especially ENet, extremely randomized trees, and NN1, outperform the simple linear models even after excluding time periods in which the machine learning models perform exceedingly well.

It may be of interest to understand what went well in 2020. One conjecture is that machine learning models can learn from difficult years (e.g., 2007 and 2008) and avoid similar missteps in 2020, which was a difficult year for real estate because of the COVID-19 crisis. To test our hypothesis, we remove the global financial crisis years (2007–2008) from our training sample and re-run our analysis. The orange bars in Figure 1 show the predictive R_{oos}^2 for OLS-2 and the three top machine learning models (ENet, extremely randomized trees, and NN1) after the global financial crisis years of 2007 and 2008 are excluded from the training set. Table E.1 in Appendix E presents the full set of results for all the models. We observe practically no change in the R_{oos}^2 for 2020 for OLS-2. However, we see large drops in the R_{oos}^2 for 2020 for the machine learning models. For example, ENet drops from an R_{oos}^2 of 13.61 percent to 3.82 percent in 2020, while the extremely randomized trees model drops from an R_{oos}^2 of 21.13 percent to 1.36 percent. This demonstrates that machine learning models have a superior ability to learn from past crises and apply the learnt lessons to crises in the future. Interestingly, although NN1 suffers a drop in predictive performance, this reduction is not as extreme as those of ENet and extremely randomized trees (its R_{oos}^2 for 2020 decreases from 23.71 to 10.49 percent). Again, this demonstrates the superiority of neural networks over less sophisticated machine learning methods.

4.4 Is the real estate market more predictable than the stock market?

In their study of the U.S. stock market, Gu et al.'s (2020) best-performing linear machine learning model has a monthly out-of-sample R^2 of 0.27 percent, while the R^2 of our best-performing linear machine learning model (ENet) is more than 10 times higher at 3.37 percent. Moreover, their best-performing regression tree has a monthly out-of-sample R^2 of 0.34 percent, while our best-performing regression tree model's R^2 of 4.52 percent is 13 times higher (for extremely randomized trees). Finally, their best-performing neural network has a monthly out-of-sample R^2 of 0.40 percent, while the R^2 of our best-performing neural network (NN1) is 5.01 percent, which is 12 times higher.

In Section 4.3, we noted that that REITs exhibit excellent out-of-sample R^2 during 2020 and 2021. However, even after excluding these breakout years, the out-of-sample predictive performance of REITs is still higher than that of stocks. For example, the last column of Table 2 shows that the R^2 of our best-performing linear machine learning model (ENet) of 1.29 percent is five times higher than the monthly out-of-sample R^2 of the best-performing linear model for stocks. Our best-performing regression tree model's R^2 of 1.23 percent (for extremely randomized trees) is four times higher than the monthly out-of-sample R^2 of the best-performing regression tree model for stocks, while the R^2 of our best-performing neural network model (NN1) of 1.46 percent is four times higher than the monthly out-of-sample R^2 of the best-performing neural network model for stocks.

These findings suggest that REIT returns are more predictable than stock returns, contrary to our initial hypothesis that the heterogeneity in the stock market would allow machine learning techniques to thrive. It also rejects the hypothesis of Nelling and Gyourko (1998), who suggest that the smaller sample of U.S. REITs than the large sample of U.S. stocks makes REITs less predictable, a reasoning that should hold true not only for simple

OLS but also for sophisticated techniques such as regression trees and neural networks.

We note some key differences between Gu et al. (2020) and our study that may make for an unfair comparison. They do not include extremely randomized trees in their repertoire of regression trees, whereas this model is our best-performing regression tree.

Further, their linear and tree models are trained on the Huber loss function instead of the default setting of the l_2 loss function that we employ in our models. Finally, their dataset spans 1953 to 2016, while ours spans 1990 to 2021. Significant differences in time periods such as the lack of high-dimensional data in the earlier decades (1950s to 1980s) and existence of more unexpected market shocks such as that in 2007 may make it more difficult for machine learning algorithms to thrive. Therefore, we re-run the analysis on the U.S. stock market from 1990 to 2021 using the same 11 machine learning methods discussed in Section 3. While Gu et al. (2020) use hyperparameter tuning, we rely on the default hyperparameter settings described in Appendix D to conduct a direct comparison between the real estate market and stock market in the same time period. The second column of Table 1 reports the R^2_{oos} for the U.S. stock market from 2006 to 2021. For ease of comparison, the third column of Table 1 reports the R^2_{oos} from Gu et al. (2020).

From 2006 to 2021, the OLS model for U.S. stocks performs better than the OLS model for U.S. REITs (-2.92 percent versus -6.89 percent). This is somewhat unsurprising, as there is more variability in the cross-section of the U.S. stock characteristics and returns than the U.S. REITs. Such a result is also expected from the findings of Nelling and Gyourko (1998) because the sample of stocks is larger than that of REITs. When the OLS model is restricted to only two or three covariates (size, value, and 12-month momentum), performance for stocks barely makes into positive territory (R^2_{oos} of 0.08 percent for OLS-2 and 0.06 percent for OLS-3), contrary to the higher positive R^2_{oos} for REITs (0.36 percent for OLS-2 and 0.31 percent for OLS-3). While beyond the scope of this study, the low R^2_{oos} of OLS-2 and OLS-3 for stocks may imply that the Fama and French (1992) three-factor model and Carhart (1997) four-factor model are no longer valid, at least empirically speaking, in the context of the U.S. stock market for 1990–2021. These two models hold up well for the U.S. REIT market. As most REIT studies today still use Kenneth French's online data library of size, value, and momentum factors derived from stock market data, the time may be apt for the creation and maintenance of real estate-specific size, value, and momentum factors derived from REIT data, which researchers can use as baseline factors.

When we move onto the linear machine learning models, real estate market predictability starts to vastly outperform stock market predictability. The linear models of LASSO and Enet generate high positive predictive R^2_{oos} for the real estate market (2.49 and 3.37 percent, respectively), while they generate R^2_{oos} of 0.21 and 0.71 percent for the stock market, respectively. When we look at the regression trees, the R^2_{oos} for the stock market move into higher positive territory, whereas the R^2_{oos} for the real estate market are still four to 10 times higher than those of the stock market.

The predictive performance for stocks in 2006–2021 is 0.28, 0.27, and 0.32 percent for NN1–NN3, respectively, lower but close to Gu et al.'s (2020) findings of 0.33, 0.39, and 0.40 percent for stocks in 1987–2016. Performance also peaks at three hidden layers (NN3) for stocks in 2006–2021, meaning that our stock analysis exhibits similarities with Gu et al. (2020) in this regard. For 2006–2021, REIT predictability based on neural networks is three to 17 times higher than stock predictability.

Our study is not the only one to report high R^2_{oos} when machine learning models are applied to return prediction. Bianchi et al. (2021) report R^2_{oos} of 5.1, 6.2, and 6.1 percent from their best linear, tree-based, and neural network models, respectively. Their best tree-based model and neural network model are similar to ours (i.e., extremely randomized trees, and NN1, respectively), whereas Gu et al. (2020) and Leippold et al. (2022) have different best-performing models in these same categories. There are major differences between the design of Bianchi et al.'s (2021) study and ours. For example, they use the 128 macroeconomic variables in Ludvigson and Ng (2009), while Gu et al. (2020) and us use the eight macroeconomic variables in Welch and Goyal (2008). Additionally, instead of feeding the predictors directly into their regression models, their excess returns are regressed on the principal components of their predictors. Their dataset is also longer, running from 1971 to 2018. Nevertheless, it is still worth asking ourselves if the predictability of REITs lies between that of stocks and bonds because REITs are a hybrid of stocks and bonds?

4.5 Predictability by industry type

Section 4.4 showed that REITs are more predictable than stocks in several ways. First, our out-of-sample performance results are more than 10 times higher than the results of Gu et al. (2020). This goes against the findings of Nelling and Gyourko (1998), who believe that REITs should be less predictable than stocks because their sample size is smaller than that of stocks. Second, to address questions that REITs are trained and tested on a different time period (1990–2021) from Gu et al.'s (2020) stocks (1953–2016), we re-run the analysis using the same time period and same default hyperparameters to allow for a direct comparison. The conclusions are unchanged: REITs are more predictable than stocks even when the latter are trained on a shorter and a more recent time period. This leads us to wonder how REITs would compare against different industries in the stock market.

Table 3 presents the out-of-sample predictive R^2 of REITs and stocks in a different way, with stocks split into 14 industries according to their two-digit SIC codes (see Figure F.1 in Appendix F). First, we find that the machine learning techniques outperform traditional regression-based strategies across all 14 industries in the stock market, consistent with the findings of Gu et al. (2020) and Leippold et al. (2022) that machine learning is superior to OLS. Second, while we would expect at least one or two industries to outperform REITs just by random chance, we find that REITs outperform every subset of stocks. Using our best linear machine learning model (ENet) as an example, REITs generate an out-of-sample R^2 of 3.37 percent, while the next three best industries come in at 1.64 percent (wholesale trade), 1.38 percent (transportation), and 1.34 percent (services).

For our best tree-based model (extremely randomized trees), REITs generate an R^2_{oos} of 4.52 percent, while the next three best industries are 3.95 percent (investment holding companies), 1.50 percent (utilities), and 1.44 percent (manufacturing). For our best neural network (NN1), REITs generate an impressive R^2_{oos} of 5.01 percent, while the next three best industries rank at 2.92 percent (investment holding companies), 2.38 percent (non-bank financial institutions), and 2.20 percent (utilities).

Tables F.2 to F.15 in Appendix F present the full out-of-sample performance results of the 14 stock sectors by year. We see that while the results are mostly stable for REITs (with the exception of the global financial crisis and COVID-19 periods), the results are more volatile for stocks when divided by subsector. For example, in some years, the random forest models generate high negative R^2 values for financial institutions, construction, chemicals, IT, transportation, utilities, retail, and services.

Interestingly, the outperformance in out-of-sample predictability for REITs in 2020 (see Section 4.3) is also observed for stocks in some sectors. Table F.1 in Appendix F shows that investment holding companies, banks, and other financial institutions perform admirably in 2020. This is perhaps because these three sectors are the most similar to REITs in the sense that they are all financial institutions. Some bricks-and-mortar industries perform well in 2020, too, namely, agriculture, mining, construction, transportation, and retail. This gives us confidence that REITs' outperformance in 2020 is not an anomaly but something validated in other industries in the stock market.

To summarize, despite dividing the U.S. stock market into more than a dozen sectors that may individually outperform REITs, the results in Table 3 and Appendix F further validate our observation that REITs are superior to stocks in terms of predictability across all stock sectors, years, and machine learning techniques.

4.6 Can industry characteristics explain the outperformance?

In this section, we examine the industry-wide characteristics of REITs and stocks to assess any discernible differences between these two asset classes. While it is beyond the scope of this study to conduct a comprehensive study that involves hundreds of industry characteristics, we take on a few usual suspects.

In some quarters, there is a belief that small firms are overlooked by investors and that their returns may be easier to predict than those of large firms. For example, Leippold et al. (2022) report that small Chinese stocks are six to 10 times more predictable than large Chinese stocks. First, Table 4 debunks the misconception that REITs are "small stocks." When modern REITs were first introduced to the market in the 1990s, the belief that REITs were small stocks may have been true. The average market cap of REITs was $380 million then as opposed to the average market cap of a stock of $910 million in the 1990s. Today, the average REIT has grown to $3.91 billion, slightly shy of the average size of a stock at $4.57 billion. In our test period of 2006–2021, some industries have average market caps much smaller than those of REITs, namely, investment holding companies ($750 million), agriculture ($348 million), construction ($1.86 billion), and wholesale ($2.40 billion). Second, to counter Leippold et al.'s (2022) observation that small Chinese companies are more predictable than large Chinese companies, the scatterplot on the left-hand side of Figure 2 shows that the three industries in the U.S. stock market that are smaller than U.S. REITs have considerably worse predictive performance than REITs. Next, we move onto stock liquidity, which has been viewed as a proxy for market efficiency. Less liquidity implies less efficiency, which may permit the persistence of predictability in returns. Table 4, which displays the average trading volume of REITs and stocks, dispels this notion. In their early days, REITs were indeed less liquid than most stocks. However, today, REITs have a higher average trading volume than the average trading volume of the stock market, especially for investment holding companies, banks, agriculture, construction, and wholesale. The scatterplot on the right-hand side of Figure 2 shows that these less liquid industries, while arguably exhibiting less market efficiency than REITs, have lower predictive performance than

REITs.

4.7 Predictability by size

Next, we analyze the differences in predictability across the size of REITs. The second column of Table 5 displays the predictability of REITs in the bottom 30th percentile by market capitalization, while the third column displays the predictability of REITs in the top 30th percentile by market capitalization. The results show that large REITs are more predictable than small REITs. The difference in size predictability is large, ranging from three to four times across the models. For instance, in the extremely randomized trees model, the R^2_{oos} for large REITs is 7.06 percent, while the R^2_{oos} for small REITs is 1.78 percent.

To test the robustness of this result, we use the models trained on large REITs to predict the returns of small REITs, and vice versa. The findings in the fourth and fifth columns of Table 5 allow us to make two interesting observations. First, the predictability of small REITs increases when prediction models trained on large REITs are applied to small REITs. The improvement is significant, ranging from 44 to 86 percent for ENet, extremely randomized trees, and NN1. This suggests that the information extracted from large REITs is applicable to small REITs. Second, the predictability of large REITs drops when prediction models trained on small REITs are used; however, the lower predictability is still higher than small REITs' predictability across the board, regardless of the model type. This suggests that the high predictability of large REITs is robust and an area that deserves further study.

4.8 Predictability by property type

In this subsection, we analyze the predictability of REITs across property types. We divide our data sample into eight types: retail, residential, office, healthcare, industrial, hotel, diversified, and others. Again, the main finding remains unchanged (Table 6). With the exception of industrial REITs, the machine learning algorithms outperform the simple linear methods across all the property types. For the industrial REITs, OLS-2 and OLS-3 provide

R^2_{oos} values of 2.35 and 2.22 percent, respectively, which are higher than the R^2_{oos} of the principal component regression, random forests, and gradient-boosted regression trees models (1.79, 0.50, and 0.24 percent, respectively) but lower than the best-performing machine learning models of ENet, extremely randomized trees, and NN1 (5.40, 3.49, and 2.47 percent, respectively).

Additionally, Table 6 shows that specialist homogeneous property types such as retail, residential, office, healthcare, and hotel perform well, with the R^2_{oos} rising above 5 percent across multiple models. Sectors that have more diversified underlying property types, such as industrial, diversified, and others, do worse.

4.9 Does predictive power depend on the size of the predictor set?

In this subsection, we create a significantly smaller set of predictors comprising a few REIT-level characteristics and their interactions with the eight macroeconomic predictors. From the original 94 × 1 vector of firm characteristics, we select the most commonly used factors in the REIT literature: size (*mve*), the book-to-market ratio (*bm*), and momentum (*mom1m, mom6m, mom12m, mom36m*). Hence, the total number of covariates in $z_{i,t}$ for the reduced predictor set is $6 \times (8 + 1) = 54$.

Table 7 compares the monthly out-of-sample predictive R^2 between the full and reduced sets of predictors for best linear model (ENet), the best tree-based model (extremely randomized trees), and the best neural network (NN1), as shown in Section 4.2.

The first conclusion from Table 7 is that the linear machine learning models and regression trees perform worse when they are constrained to using a smaller set of predictors, with the R^2 dropping from 3.37 to 0.35 percent for ENet and from 4.52 to 2.86 percent for extremely randomized trees. Again, this is unsurprising, as the regularization technique of the linear machine learning models and feature selection ability of the regression trees are well suited to dealing with high-dimensional data, so more predictors are preferred over fewer predictors. That said, REIT predictability still outperforms stock predictability by three times.

The second conclusion from Table 7 is somewhat surprising and deserves further study. When faced with a reduced set of predictors, the predictive performance of NN1 does not suffer. This suggests that neural networks may be superior to other machine learning techniques in teasing out the non-linearities and complex interactions between predictors, even when faced with a much smaller set of variables.

4.10 Which predictors matter?

Given the stark differences in predictability between stocks and REITs, we investigate whether certain predictors are more important than others within the real estate domain to shed light on why machine learning models, when applied to predict real estate returns, perform admirably compared with their prediction of stock returns.

To identify predictors that have an important influence on the cross-section of expected REIT returns, while simultaneously controlling for other predictors in the dataset, we rank all the predictors by variable importance, which we denote as VI_j for the jth predictor. We calculate the reduction in R^2 by setting all the values of predictor j within each training sample to zero, while holding the remaining predictors fixed. We then average them into a single importance measure for each predictor.

We first explore the variable importance of the eight macroeconomic variables for all the prediction models. Here, variable importance is defined as in Gu et al. (2020) and Leippold et al. (2022), namely, for a specific model. Table 8 reports the relative variable importance of our eight macroeconomic variables: the dividend-to-price ratio (*dp*), earnings-to-price ratio (*ep*), book-to-market ratio (*bm*), net equity expansion (*ntis*), Treasury bill rate (*tbl*), term spread (*tms*), default spread (*dfy*), and stock variance (*svar*). Figure 3 compares the macroeconomic variables across the models.

In all the models, market volatility is a critical predictor, whereas the aggregate book-to- market ratio has little role in most models for the REIT market. This is in sharp contrast to the findings of Gu et al. (2020), who show that the aggregate book-to-market ratio is the most important predictor for the stock market and that market volatility (*svar*) is the least important. Guo (2006) shows that stock variance is a proxy for market risk and can drive out most variables such as the dividend yield, default premium, and term premium when predicting stock returns from 1952 to 2002. While outside the scope of this study, uncovering the theoretical underpinnings of why Guo's (2006) findings apply so strongly to the REIT market would be an exciting future challenge.

The next most important macroeconomic predictor for most models is the default spread, which is defined as the difference between BAA- and AAA-rated corporate bond yields. This is commonly viewed as another proxy for market risk. Fama and French (1989) find that the default spread, which tends to be higher when business conditions are weak, can explain the returns of both bond and stock portfolios. A possible explanation why *dfy* performs so well in our study is that REITs are typically regarded as a hybrid of stocks and bonds in terms of their short-run return and risk exposure (e.g., Karolyi & Sanders, 1998; Ling & Naranjo, 1999; Peterson & Hsieh, 1997).

The two most important time series predictors, *svar* and *dfy*, for our strongest machine learning models (neural networks) are practically ignored by Gu et al.'s (2020) best-performing machine learning models (also neural networks). For reference, Appendix G provides Gu et al.'s (2020) relative variable importance for the macroeconomic predictors. As an example, the variable importance of *svar* and *dfy* in the U.S. stock market for NN1 is 0.57 and 0.09 percent, respectively, whereas the variable importance of the same two predictors in the U.S. REIT market is 65.5 and 27.0 percent, respectively. This finding suggests that the real estate market and stock market are sensitive to very different macroeconomic environments, an area worthy of further study.

Not all of the REIT characteristics are equally important in predicting REIT returns and their importance depends greatly on the prediction model. Figure 4 reports the variable importance of the top 20 firm-level REIT characteristics for the best linear method, regression tree, and neural network. The variable importance within each model is normalized to one to allow us to interpret the relative importance of each model. We also include the baseline OLS model for comparison purposes.

Appendix A provides the details of all the firm-level characteristics.

Figure 5 reports the overall rankings of the characteristics of all the models. We rank the importance of each characteristic for each method and then sum their ranks. The characteristics are ordered so that the highest total ranks are at the top and the lowest ranking characteristics are at the bottom. The color gradient within each column shows the model-specific ranking of the characteristics from the least to most influential (white to dark blue).

Figure 4 shows clearly that OLS has selected two unusual characteristics (*rd_sale*, *rd_mve*) as its most important predictors, with the beta-related variables (*beta*, *betasq*) ranking third and fourth. The R&D-to-sales and R&D-to-market capitalization ratios are firm-level characteristics that do not seem to have any direct relevance to REITs, which may explain why the OLS model produces a predictive R^2_{oos} of -6.89 percent.

As shown in Figure 5, the machine learning models generally agree that *securedind* is the most influential REIT-level predictor, whereas the top predictor for the U.S. stock market is *mom1m* (Gu et al., 2020). This implies that risk measures (*securedind*) contain predictive ability for REITs, while short-term price trends (*mom1m*) contain predictive ability for stocks. Using a large sample of publicly traded firms from 1985–2012, Valta (2016) presents evidence that expected stock returns are higher for firms that have a large proportion of secured debt (*securedind*) or convertible debt (*convind*). While convertible debt is not commonly issued in the REIT industry, it is worth further study to investigate the link between secured debt and REIT returns given its high predictive ability.

Like Gu et al. (2020), the REIT-level predictors can be grouped into four major categories. The first is based on recent price trends, which occupy six of the top 15 variables in Figure 5 (*mom12m*, *mom1m*, *mom6m*, *mom36m*, *indmom*, *maxret*).[6] This agrees with Gu et al. (2020), who find that price trends comprise the largest group of top predictors for U.S.

---

[6] Appendix A provides the full description of these predictors and their references.

stocks. The next group is based on risk measures, which account for five of the top 15 importance variables (*securedind*, *beta*, *retvol*, *betasq*, *lev*). Valuation ratios and fundamental signals constitute the third influential group (*dy*, *sp*, *roic*). The last group comprises liquidity variables, which has only one variable of the top 15 (*baspread*).

It is worthwhile taking a deeper dive to understand the drivers behind the high R^2_{oos}s exhibited by the individual machine learning models. As shown in Figure 4, ENet puts most of its weight (98.7 percent) behind *secureind*. Heavily penalizing 93 of the 94 characteristics to mainly rely on a single predictor to forecast returns does not make for a robust asset pricing model. We see hints of this behavior in Table 8, where ENet puts 98.5 percent of its weight behind a single macroeconomic predictor (*svar*), while non-linear models such as extremely randomized trees and NN1 draw predictive information from a more diversified set of predictors. This extreme behavior is also exhibited by LASSO, another linear machine learning model. The inability of linear machine learning models to accommodate non-linear relationships in the real estate domain, which results in them penalizing a large swath of informative predictors, could explain why such models do not predict returns as well as regression trees and neural networks do.

Appendix H shows the trend of the top 10 predictors over the entire test period for the OLS, ENet, extremely randomized trees, and NN1 models. We observe that all the models each stay with the same top predictors from 2006 to 2021, which gives us comfort that the machine learning models are relatively stable over time. The only difference is in the relative weights assigned to these predictors. Using *securedind* as an example, ENet shows supreme confidence in using it as its top predictor early on, giving it a relative weight of more than 90 percent in 2006 and keeping it a high level until 2021 (see Table H.2). For extremely randomized trees, the weight given to *securedind* takes a few years to stabilize, staying constant at 45–46 percent only from 2010 onward (see Table H.3). NN1 takes an even longer time to settle, landing on a relative weight of 80–90 percent for *securedind* from 2015 (see Table H.4). This suggests that a more sophisticated machine learning model such as NN1 needs a longer dataset to stabilize its predictor weights than a simpler model such as ENet needs. However, while a more sophisticated model waits for its weights to stabilize over time, this does not come at the expense of predictive performance. NN1 consistently outperforms extremely randomized trees, which in turn outperform ENet.

5 Portfolio analysis

5.1 Prespecified portfolio forecasts

Thus far, we have analyzed the predictability of individual REIT returns. Next, we examine performance at the aggregate portfolio level for several reasons. As all of our models are optimized for firm-level forecasts, portfolio-level forecasts provide an additional evaluation of our machine learning models and their robustness. Aggregate portfolios also tend to be of economic interest because they are commonly held by individual investors through retirement savings funds, mutual funds, and exchange-traded funds. The distribution of portfolio-level monthly returns is sensitive to dependence among individual REIT returns and whether a good REIT-level prediction model can produce accurate aggregate-level forecasts is unknown.

We build bottom-up forecasts by aggregating the individual REIT returns into portfolios. Given the weight of the *i*th REIT in portfolio *p*, which we denote as $w_{i,t}^{p}$, and an out-of-sample forecast of the *i*th REIT, which we denote as $\hat{r}_{i,t+1}$, we construct the aggregate portfolio return forecast as

$$\hat{r}_{i,t+1}^{p} = \sum_{i=1}^{n} w_{i,t}^{p} \times \hat{r}_{i,t+1} \quad (5)$$

We form bottom-up forecasts for both the value-weighted and the equally weighted portfolios of REITs. Table 9 reports the monthly out-of-sample R^2 over our 15-year testing sample. We see a marked improvement in predictability. The monthly R_{oos}^2 of our best linear, tree, and neural network models experience a two-to-threefold jump when forecasting at the aggregate level, with ENet, extremely randomized trees, and NN1 reporting R_{oos}^2 values of 10.79, 11.46, and 11.86 percent for the value-weighted portfolios, respectively. We observe qualitatively similar results for the equally weighted portfolios.

Other studies have shown the difficulty in producing high out-of-sample R^2_{oos} at the aggregate portfolio level. For example, the macroeconomic predictors used by Welch and Goyal (2008) fail to produce a positive out-of-sample R^2 for the stock market (despite producing excellent in-sample R^2 for the same predictors). Using a partial least squares machine learning model, Kelly and Pruitt (2013) find an out-of-sample R^2 of just 1 percent for the aggregate market index. Gu et al.'s (2020) out-of-sample R^2 for the generalized linear model is 0.71 percent, which reaches as high as 1.40 to 1.80 percent for their best tree-based and neural network models. Again, REITs outperform. The strength of our results for U.S. REITs suggests that machine learning has a positive role to play in the investment and portfolio construction of U.S. REITs, which have a combined equity market capitalization of $1.3 trillion and are owned by 150 million Americans.[7]

5.2 Machine learning portfolios

Next, we create new sets of portfolios to directly exploit the machine learning forecasts. At the end of each month, we calculate the 1-month-ahead out-of-sample REIT return predictions for each machine learning model. We then sort the REITs into two groups based on the breakpoints for the bottom 30 and top 30 percent of each model's forecasted returns. We reconstitute machine learning portfolios each month using value weights. We construct a long-only portfolio that only holds REITs with the highest expected returns (top 30 percent) and a zero-net-investment portfolio that buys REITs with the highest expected returns (top 30 percent) and sells REITs with the lowest expected returns (bottom 30 percent). These tests provide an additional evaluation of our machine learning models and their robustness. Table 10 reports the out-of-sample performance for the value-weighted long-only and value-weighted long-short portfolios for our best-performing linear, tree, and neural network machine learning models (i.e., ENet, extremely randomized trees, and NN1, respectively). In

[7] Source: Nareit, as of 1 April 2023. https://www.reit.com/data-research/data/reits-numbers.

Panel A, the long-only machine learning portfolios are benchmarked against a value-weighted index of all U.S. REITs. In Panel B, the long-short machine learning portfolios are benchmarked against a long-short portfolio based on predictions from the best-performing OLS model in Table 1 (i.e. the Fama–French-inspired OLS-2 model). Figure 6 charts the cumulative performance of the benchmark portfolios and machine learning portfolios constructed by the top models.

Once again, the tree-based models and neural networks dominate in terms of average returns and Sharpe ratios. Interestingly, the long-short portfolios of the neural networks exhibit a positive skew in returns, an attribute favored by investors. Most equity market portfolios exhibit negative skewness, which partly explains why equity risk premiums exist, as investors are more averse to downside than upside risks. Investors do not complain when markets rise. The long-short portfolios also exhibit negative correlations to the aggregate market index of all U.S. REITs, another attribute favored by investors. This means that long-short portfolios generated by machine learning models have the potential to be effective investment overlays for a buy-and-hold strategy, namely, by lowering overall portfolio volatility, while increasing expected returns.

However, although the overall performance statistics look good, our machine learning portfolios seem to suffer from a drop in profitability after 2016. While there are insufficient data points to make the definitive conclusive that the machine learning portfolios are no longer profitable (a portfolio with a Sharpe ratio of 0.3 to 0.6 may be expected to experience a period of investment drought), one cannot help but note that Nvidia started tweaking GPUs to handle specific A.I. calculations in 2017. That same year, Nvidia also began selling complete computers to carry out A.I. tasks more efficiently than its past strategy of selling chips or circuit boards for other companies' systems.[8] Therefore, the increased accessibility of GPUs to

[8] https://www.nytimes.com/2023/08/21/technology/nvidia-ai-chips-gpu.html.

the public could be a contributing factor to the decreased profitability of the machine learning models after 2016.

We also note the lackluster investment performance of ENet despite its outperformance in the out-of-sample R^2-based tests in Tables 1 and 9. As shown in Panel B of Table 10, the long-short performance of ENet is negative, returning an average of -0.15 percent per month, while extremely randomized trees and NN1 generate average monthly returns of 0.41 and 0.35 percent, respectively. On a risk-adjusted basis, ENet generates a negative annualized Sharpe ratio of -0.11, while the naive long-short portfolio based on Fama and French's size and value factors (OLS-2) generates a negative Sharpe ratio of -0.06. On the contrary, extremely randomized trees and NN1 are squarely in positive territory, with Sharpe ratios of 0.35 and 0.29, respectively. This is the first sign of a crack in the robustness of the linear machine learning models, as they fail to take advantage of the non-linearities contrary to extremely randomized trees and NN1. Table 11 shows that the standard deviation of the predicted monthly returns for ENet is 1.04, while the standard deviations for extremely randomized trees and NN1 are 30–100 percent higher at 1.36 and 2.14, respectively. The predicted returns for ENet range from -12.19 to 1.64 percent, whereas the ranges for extremely randomized trees and NN1 are from -25.32 to 4.77 percent and from -53.50 and 2.06 percent, respectively. The inability of ENet to be bolder in its return predictions is the likely cause of its failure to segregate high-performing REITs (top 30 percent) from low-performing REITs (bottom 30 percent) in the long-short strategy. The real world is considerably more colorful and complicated. The first column of Table 11 shows that the actual standard deviation of monthly returns is 11.83, with a wide range of returns from -90.93 to 290.27 percent.

Value-weighted strategies are important, as they reflect market reality by considering the market capitalization of each REIT, and larger REITs do have a larger impact on the real estate market than smaller ones. Nevertheless, it is also useful to study equally weighted strategies in our analysis, as our statistical objective functions minimize equally weighted forecasting errors. Table I.1

in Appendix I reports the performance of the machine learning portfolios using an equal-weight construction and Figure I.1 in the same appendix charts the cumulative performance of the equally weighted machine learning portfolios constructed by the top models. The results are qualitatively unchanged: the regression trees and neural networks outperform, while the linear machine models underperform.

We next consider a meta-strategy that takes advantage of our main finding thus far that regression trees and neural networks seem to be good prediction models. We construct an average of the portfolios derived from all the regression trees and neural networks in our toolkit, leading to a grand ensemble of eight non-linear methods comprising random forests, gradient-boosted regression trees, extremely randomized trees, and NN1–NN5. The last rows of Panels A and B in Table 10 show that such an ensemble of non-linear models delivers the best risk-adjusted performance in both the long-only and the long-short portfolios of REITs, with annualized Sharpe ratios of 0.60 and 0.50, respectively, higher than any single non-linear method on its own. For the long-short portfolio, the ensemble generates the lowest maximum drawdown and lowest maximum one-month loss, better than any single non-linear method on its own. This illustrates the economic potential of using machine learning models within the asset pricing and investment fields.

5.3 Mean-variance portfolios

Allen et al. (2019) challenge the academic consensus that the mean-variance portfolio fails to outperform passive equal-weighted approaches. The mean–variance approach seeks to overweight assets with low correlations, high expected returns, and low relative variance. One common criticism is that return and covariance forecast errors are magnified in the estimation of portfolio weights, which tend to lead to poor out-of-sample performance. Indeed, in Markowitz's (1952) original formulation, he states that "we must have procedures for finding reasonable μ_i and σ_{ij} . These procedures, I believe, should combine statistical techniques and the judgment of practical men." He reiterates this view 58 years later: "judgment plays an essential role in the proper application of risk–return analysis for individual and institutional portfolios. For example, the estimates of mean, variance, and covariance of a mean–variance analysis should be forward-looking rather than purely historical" (Markowitz, 2010). In short, if one has good predictive ability, use it. This is consistent with Markowitz's view of mean-variance optimization. If not, equal weighting or value weighting a portfolio may be preferable.

We advance Allen et al.'s (2019) work by incorporating the machine learning forecasts into mean-variance optimization. We use the initial training sample of 192 months (1990–2005) to estimate the historical mean and covariance matrix of REIT returns, which are used to form our sample-based mean-variance portfolio. Using the same covariance matrix, we form machine learning mean-variance portfolios by replacing the historical means with the expected returns from our best-performing machine learning models. At the end of each month, we add a further month of data into the training sample and update the expected mean and covariance matrix of the REIT returns. The entire out-of-sample test period is 192 months (2006–2021). This strategy ignores the possibility of estimation error in the covariance matrix. However, by applying the same covariance matrix to the machine learning portfolios, it allows us to find whether an investor endowed with machine learning techniques benefits from a mean-variance approach despite such estimation errors.

Table 12 reports the out-of-sample performance for our machine learning mean-variance models compared with the traditional sample-based mean-variance portfolio and naive $1/N$ portfolio. The naive strategy involves holding a portfolio weight of $1/N$ in each of the N REITs for each time period, and this generates a Sharpe ratio of 0.53. In Panel A, the mean-variance portfolios are constrained to long-only positions to allow for a direct comparison with the naive $1/N$ portfolio, which is long-only. Unsurprisingly, the traditional sample-based mean-variance portfolio, with its Sharpe ratio of 0.48, cannot overcome the inherent estimation errors and outperform the naive $1/N$ portfolio. Similarly, the ENet portfolio, with its Sharpe ratio of 0.48, also cannot outperform the naive portfolio in risk-adjusted terms. Such underperformance was noted in Section 5.2. Extremely randomized trees, NN1, and the ensemble of all the non-linear machine learning portfolios outperform the naive 1/*N* portfolio, with Sharpe ratios of 0.65, 0.58, and 0.61, respectively. In Panel B, we allow the mean-variance portfolios to take long-short positions to explore whether they make for effective overlays on top of a long-only strategy. We see the sample-based mean-variance portfolio struggling in this instance. It has a Sharpe ratio of 0.18 and a correlation with the naive 1/*N* portfolio of 0.37. Conversely, extremely randomized trees, NN1, and the non-linear ensemble generate much higher Sharpe ratios of 0.56, 0.46, and 0.54, respectively, while exhibiting very low correlations with the 1/*N* portfolio of 0.01, 0.00, and 0.07. Once again, the machine learning techniques prove their worth in portfolio management.

5.4 Time decay analysis

In this subsection, we examine the decay of return predictability over time. Table 13 shows the out-of-sample R^2 of the predicted returns of our top machine learning models trained on monthly returns compared with the actual realized returns over the next month, next three months, next six months, and next 12 months. It is clear that nothing good lasts forever. The time decay is brutal, with the R^2s turning negative for ENet and NN1 after six months. For extremely randomized trees, the R^2 is almost zero after 12 months. This clearly shows that information decay sets in quickly for our prediction models. It could also suggest that our top machine learning models are ruthless in using the 800+ predictors to achieve their objective function, which focuses on predicting returns one month out at a time. These models are unconcerned if they are bad at predicting returns further out in time, say three or six months out. For instance, NN1 may weigh short-term signals such as 1-month momentum heavily if the objective function is to predict returns one month out, but choose to weigh fundamental signals such as the book-to-price ratio and leverage ratio heavily if its objective function is to predict one year out.

To test our hypothesis, we re-run our analysis with varying objective functions. Instead of using monthly returns in the training set, we switch to using quarterly, biannual, and annual returns in the training sets and the models are required to predict returns for the next three, six, and 12 months, respectively. True enough, when the top models' objective function is to predict returns over the next 12 months, they can change tack and do well, just as they can closely predict the next monthly return if their objective function states so (Table 14). While extremely randomized trees seem to struggle slightly with quarterly and biannual predictions, this may be solved with some hyperparameter tuning, which is outside the scope of this study. While it is interesting to observe the results in Table 13, the more important lesson learnt here is that machine learning models are terrible at predicting something for which they are not trained.

6 Conclusion

Machine learning algorithms can improve our understanding of real estate returns, both in economic research and in portfolio management. This study adds to the burgeoning evidence that machine learning methods can be successfully applied to markets that are fundamentally different from the mainstream U.S. stock market. Overall, our findings demonstrate that machine learning can help improve the empirical understanding of real estate asset pricing. Neural networks and, to a lesser extent, regression trees are the best-performing methods. Shallow learning outperforms deep learning in our case, which differs from the typical conclusions in other data science fields. We also find that the predictive performance of REIT returns is generally higher than general stock returns in all periods; although the machine learning techniques failed to predict the 2007 subprime crisis, they learned from it and avoided repeating such mistakes when faced with subsequent market turbulence (e.g., the 2020 COVID-19 crisis). We observe that the returns of large REITs are more predictable than those of small REITs, while the returns of specialist REITs are more predictable than the performance of REITs with diversified underlying properties.

Contrary to Li and Wang (1995) and Nelling and Gyourko (1998), who find that REIT returns are not more predictable than stock returns, we find that real estate returns are *significantly more predictable* than stock returns, with predictability up to 12 times higher. When we subdivide the stock market into 14 industries by their SIC code, hoping that one or two industries may outperform REITs, we find that none are able to.

Machine learning predictions hold up well, not only at the individual REIT level, but also for portfolio-level returns, with the R^2_{oos} exceeding 10 percent for the best linear, tree, and neural network models. The evidence for economic gains from machine learning forecasts is likewise impressive, with higher Sharpe ratios and t-statistics. We wish to highlight that none of the machine learning models used in our study undergo hyperparameter tuning, unlike those employed by Gu et al. (2020), Bianchi et al. (2021), and Leippold et al. (2022). With more powerful processing capabilities to conduct hyperparameter optimization, we expect the performance results for the REIT market to be better than the performance found in this study.

Finally, the relative decline in predictability across all models starting around 2016 could be an outcome of the perpetual arms race in data, models and trade execution speed that is at the core of market efficiency. Our machine learning portfolios display lower profitability from around 2016 onward, the time when Nvidia started making GPUs more accessible to the public. A similar pattern in predictive power appears in Gu et al's (2020) general stock market analysis: The cumulative returns on portfolios based on the dominating machine learning model (NN4) shown in their Figure 9 appears to flatten in the 2010s. We expect the quantitative arms race to intensify further as data and algorithms improve further, and we are looking forward to more updates from academic studies soon.

Acknowledgments

We are grateful to Brent Ambrose, Colin Lizieri, Philip Kalikman, Rongjie Zhang, Kelvin Wong, Daniel Ruf, Chongyu Wang, Wayne Wan, Derek Ho, Norm Miller, Arpit Gupta, Stijn van Nieuwerburgh, Gianluca Marcato, and the participants of the 2023 AREUEA International Meeting, 2023 Real Estate Finance and Investment Symposium at the University of Hong Kong, and 2024 Weimer School Session for their invaluable comments and suggestions. We also thank Dacheng Xiu for sharing the empirical data on the stock-level characteristics and Amit Goyal for sharing his database of macroeconomic predictors. Finally, we are indebted to Daniel McMillen for asking constructive but hard "why" questions.

References

Allen, D., Lizieri, C., & Satchell, S. (2019). In defense of portfolio optimization: What if we can forecast? *Financial Analysts Journal, 75*(3), 20–38. https://doi.org/10.1080/0015198X.2019.1600958

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance, 61*(1), 259–299. https://doi.org/10.1111/j.1540-6261.2006.00836.x

Asness, C. S., Porter, R. B., & Stevens, R. L. (2000). Predicting stock returns using industry-relative firm characteristics. *Working Paper, AQR.* https://doi.org/10.2139/ssrn.213872

Baldominos, A., Blanco, I., Moreno, A. J., Iturrarte, R., Bernárdez, Ó., & Afonso, C. (2018). Identifying real estate opportunities using machine learning. *Applied Sciences, 8*(11), 2321. https://doi.org/10.3390/app8112321

Bazdresch, S., Belo, F., & Lin, X. (2014). Labor hiring, investment, and stock return predictability in the cross-section. *Journal of Political Economy, 122*(1), 129–177. https://doi.org/10.1086/674549

Bianchi, D., Buchner, M., & Tamoni, A. (2021). Bond risk premiums with machine learning. *The Review of Financial Studies, 34*(2), 1046–1089. https://doi.org/10.1093/rfs/hhaa062

Blau, B. M., Hill, M. D., & Wang, H. (2011). REIT short sales and return predictability. *Journal of Real Estate Finance and Economics, 42*(4), 481–503. https://doi.org/10.1007/s11146-009-9196-9

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Bühlmann, P., & Hothorn, T. (2007). Boosting algorithms: Regularization, prediction and model fitting. *Statistical Science, 22*(4), 477–505. https://doi.org/10.1214/07-STS242

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Chordia, T., Subrahmanyam, A., & Anshuman, V. R. (2001). Trading activity and expected stock returns. *Journal of Financial Economics, 59*(1), 3–32. https://doi.org/10.1016/S0304-405X(00)00080-5

Chui, A., Titman, S., & Wei, K. (2003). The cross-section of expected REIT returns. *Real Estate Economics, 31*(3), 451–479. https://doi.org/10.1111/1540-6229.00073

Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems, 2*(4), 303–314. https://doi.org/10.1007/BF02551274

Elkind, D., Kaminski, K., Lo, A., Siah, K. W., & Wong, C. H. (2022). When do investors freak out? Machine learning predictions of panic selling. *Journal of Financial Data Science, 4*(1), 11–39. https://doi.org/10.3905/jfds.2021.1.085

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*(1), 23–49. https://doi.org/10.1016/0304-405X(89)90095-0

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*(2), 427–465. https://doi.org/10.2307/2329112

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of*

Statistics, 29(5), 1189–1232. https://doi.org/10.1214/aos/1013203451

Giacomini, E., Ling, D. C., & Naranjo, A. (2015). Leverage and returns: A cross-country analysis of public real estate markets. *Journal of Real Estate Finance and Economics, 51*(2), 125–159. https://doi.org/10.1007/s11146-014-9489-5

Green, J., Hand, J., & Zhang, F. (2017). The characteristics that provide independent information about average U.S. monthly stock returns. *The Review of Financial Studies, 30*(12), 4389–4436. https://doi.org/10.1093/rfs/hhx019

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies, 33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning, 63*(1), 3–42. https://doi.org/10.1007/s10994-006-6226-1

Guo, H. (2006). On the out-of-sample predictability of stock market returns. *Journal of Business, 79*(2), 645–670. https://doi.org/10.1086/499134

Hoerl, A., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*(1), 55–67. https://doi.org/10.1080/00401706.1970.10488635

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks, 2*(5), 359–366. https://doi.org/10.1016/0893-6080(89)90020-8

Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *International Conference on Machine Learning,* 448–456.

Karolyi, G. A., & Sanders, A. (1998). The variation of economic risk premiums in real estate returns. *Journal of Real Estate Finance and Economics, 17*(3), 245–262. https://doi.org/10.1023/A:1007776907309

Kelly, B., & Pruitt, S. (2013). Market expectations in the cross-section of present values. *Journal of Finance, 68*(5), 1721–1756. https://doi.org/10.1111/jofi.12060

Kelly, B., Pruitt, S., & Su, Y. (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics, 134*(3), 501–524. https://doi.org/10.1016/j.jfineco.2019.05.001

Kingma, D., & Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint,* arXiv:1412.6980.

Leippold, M., Wang, Q., & Zhou, W. (2022). Machine learning in the Chinese stock market. *Journal of Financial Economics, 145*(2), 64–82. https://doi.org/10.1016/j.jfineco.2021.08.017

Li, Y., & Wang, K. (1995). The predictability of REIT returns and market segmentation. *Journal of Real Estate Research, 10*(4), 471–482. https://doi.org/10.1080/10835547.1995.12090800

Lin, C. Y., Rahman, H., & Yung, K. (2009). Investor sentiment and REIT returns. *Journal of Real Estate Finance and Economics, 39*(4), 450–471. https://doi.org/10.1007/s11146-008-9115-5

Ling, D. C., & Naranjo, A. (1999). The integration of commercial real estate markets and stock markets. *Real Estate Economics, 27*(3), 483–515. https://doi.org/10.1111/1540-6229.00781

Ling, D. C., Naranjo, A., & Ryngaert, M. D. (2000). The predictability of equity REIT returns: Time variation and economic significance. *Journal of Real Estate Finance and Economics, 20*(2), 117–136. https://doi.org/10.1023/A:1007821103728

Lindenthal, T., & Johnson, E. B. (2021). Machine learning, architectural styles and property values. *Journal of Real Estate Finance and Economics*. https://doi.org/10.1007/s11146-021-09845-1

Liu, C. H., & Mei, J. (1992). The predictability of returns on equity REITs and their co-movement with other assets. *Journal of Real Estate Finance and Economics, 5*(4), 401–418. https://doi.org/10.1007/BF00174808

Ludvigson, S. C., & Ng, S. (2009). Macro factors in bond risk premia. *Review of Financial Studies, 22*(12), 5027–5067. https://doi.org/10.1093/rfs/hhp081

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*(1), 77–91. https://doi.org/10.2307/2975974

Markowitz, H. M. (2010). Portfolio theory: As I still see it. *Annual Review of Financial Economics, 2,* 1–23. https://doi.org/10.1146/annurev-financial-011110-134602

Masters, T. (1993). *Practical neural network recipes in C++*. New York, NY: Academic Press.

Moritz, B., & Zimmermann, T. (2016). Tree-based conditional portfolio sorts: The relation between past and future stock returns. *Working Paper, Ludwig Maximilian University of Munich.* https://doi.org/10.2139/ssrn.2740751

Nelling, E., & Gyourko, J. (1998). The predictability of equity REIT returns. *Journal of Real Estate Research, 16*(3), 251–268. https://doi.org/10.1080/10835547.1998.12090958

Ooi, J., Wang, J., & Webb, J. (2009). Idiosyncratic risk and REIT returns. *Journal of Real Estate Finance and Economics, 38*(4), 420–442. https://doi.org/10.1007/s11146-007-9091-1

Pai, P. F., & Wang, W. C. (2020). Using machine learning models and actual transaction data for predicting real estate prices. *Applied Sciences, 10*(17), 5832. https://doi.org/10.3390/app10175832

Palazzo, B. (2012). Cash holdings, risk, and expected returns. *Journal of Financial Economics, 104*(1), 162–185. https://doi.org/10.1016/j.jfineco.2011.12.009

Peterson, J., & Hsieh, C. (1997). Do common risk factors in the returns on stocks and bonds explain return on REITs? *Real Estate Economics, 25*(2), 321–345. https://doi.org/10.1111/1540-6229.00717

Rapach, D. E., Strauss, J. K., & Zhou, G. (2013). International stock return predictability: What is the role of the United States? *Journal of Finance, 68*(4), 1633–1662. https://doi.org/10.1111/jofi.12041

Serrano, C., & Hoesli, M. (2010). Are securitized real estate returns more predictable than stock returns? *Journal of Real Estate Finance and Economics, 41*(2), 170–192. https://doi.org/10.1007/s11146-008-9162-y

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research, 15*(1), 1929–1958.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 58*(1), 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Valta, P. (2016). Strategic default, debt structure, and stock returns. *Journal of Financial and Quantitative Analysis, 51*(1), 197–229. https://doi.org/10.1017/S002210901600003X

Viriato, J. C. (2019). AI and machine learning in real estate investment. *The Journal of Portfolio Management, 45*(7), 43–54. https://doi.org/10.3905/jpm.2019.45.7.043

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies, 21*(4), 1455–1508. https://doi.org/10.1093/rfs/hhm014

Zhou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67*(2), 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

Tables

Table 1: Monthly out-of-sample REIT-level prediction performance (percentage R^2_{oos})

	REITs	*Stocks*	*Stocks (Gu et al.)*	*Bonds (Bianchi et al.)*	*Stocks (Leippold et al.)*
OLS	-6.89	-2.92	-3.46	N.A.	0.81
OLS-2	0.36	0.08	N.A.	N.A.	N.A.
OLS-3	0.31	0.06	0.16	N.A.	0.77
LASSO	2.49	0.21	N.A.	5.10	1.43
ENet	3.37	0.71	0.11	4.80	1.42
PCR	0.28	0.07	0.26	-4.90	N.A.
RF	2.71	0.54	0.33	3.90	2.44
GBRT	2.70	0.25	0.34	-1.80	2.71
ERT	4.52	1.03	N.A.	6.20	N.A.
NN1	5.01	0.28	0.33	6.10	2.07
NN2	2.09	0.27	0.39	1.40	2.04
NN3	1.02	0.32	0.40	N.A.	2.28
NN4	0.75	0.20	0.39	N.A.	2.49
NN5	0.79	0.00	0.36	N.A.	2.58
Test Period	2006–2021	2006–2021	1987–2016	1990–2018	2012–2020

Notes: The first two columns report monthly $R^2_{oos}$for the entire panel of REITs and stocks using OLS with all variables (OLS), OLS using only size and book-to-market (OLS-2), OLS using only size, book-to-market, and 12- month momentum (OLS-3), least absolute shrinkage and selection operator (LASSO), elastic net (ENet), principal component regression (PCR), random forest (RF), gradient boosted regression trees (GBRT), extremely randomized trees (ERT), and neural networks with one to five layers (NN1–NN5). The third column displays the corresponding prediction performance for the

U.S. stock market as presented in Gu et al. (2020). The fourth columns displays the out-of-sample performance for the U.S. bond market as presented in the corrigendum for Bianchi et al. (2021), for bonds with 13-24 months of maturity, and with forward rates and macroeconomic variables as forecasting variables. The last column displays the corresponding prediction performance for the Chinese stock market as presented in Leippold et al. (2022). All the numbers are expressed as a percentage.

Table 2: Monthly predictive R^2_{oos}, by year

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-42.14	-15.42	-10.29	-7.72	-2.12	-0.44	-3.14	-22.08	-37.23
OLS-2	3.91	-4.74	-2.00	0.73	2.15	-1.18	3.84	1.37	3.26
OLS-3	4.29	-4.16	-1.58	0.28	2.23	-1.21	3.36	1.81	3.03
LASSO	4.27	-4.44	-1.96	0.36	3.45	1.68	4.55	1.29	3.38
ENet	4.41	-4.27	2.15	0.11	3.64	1.80	4.64	1.27	3.38
PCR	4.16	-4.49	-1.80	0.48	2.02	-1.28	3.73	1.36	3.61
RF	4.73	-9.97	2.43	0.62	2.57	-0.87	3.71	-0.04	4.36
GBRT	5.91	-7.80	5.73	-0.34	0.52	3.52	2.34	0.47	4.93
ERT	4.49	-5.18	3.89	0.09	2.48	0.07	4.14	1.25	4.00
NN1	2.68	-2.67	1.56	1.04	3.78	0.74	4.51	1.29	2.16
NN2	3.08	-3.09	-1.59	1.07	2.61	-0.90	4.43	1.19	1.88
NN3	3.18	-2.95	-1.73	1.06	2.58	-1.09	4.38	1.22	2.21
NN4	2.98	-2.94	-1.67	1.07	2.51	-1.14	4.39	1.17	1.98
NN5	2.99	-2.93	-1.69	1.05	2.51	-1.11	4.36	1.17	1.84

	2015	2016	2017	2018	2019	2020	2021	All Years	All ex '20-21
OLS	-7.30	-11.72	-6.04	0.75	-10.12	7.82	1.19	-6.89	-10.10
OLS-2	-2.01	2.45	1.10	-3.37	3.82	-0.04	3.64	0.36	0.25
OLS-3	-1.85	2.03	1.05	-3.15	3.69	-0.15	3.54	0.31	0.22
LASSO	-1.65	2.93	0.82	-1.28	3.81	11.69	4.21	2.49	0.56

ENet	-1.64	2.99	0.75	-0.93	3.78	13.61	4.22	3.37	1.29
PCR	-1.81	2.23	0.75	-3.74	3.64	-0.23	3.64	0.28	0.19
RF	-2.70	1.96	1.04	-6.34	3.84	12.94	4.25	2.71	0.59
GBRT	-3.80	1.73	0.84	-3.42	-1.50	11.09	3.61	2.70	0.98
ERT	-2.31	2.57	0.96	-4.15	4.13	21.13	4.23	4.52	1.23
NN1	0.05	2.38	-0.30	2.30	2.30	23.71	1.85	5.01	1.46
NN2	-0.69	2.11	1.05	0.89	1.94	9.95	0.34	2.09	0.62
NN3	-0.90	2.08	1.10	-2.78	2.41	3.93	0.50	1.02	0.48
NN4	-0.12	1.88	0.03	-0.72	1.41	2.49	-0.73	0.75	0.49
NN5	-0.60	1.74	1.01	-1.19	0.76	2.34	1.43	0.79	0.44

Notes: This table reports monthly R^2_{oos} for the entire panel of REITs, by calendar year. All the numbers are expressed as a percentage.

Table 3: Monthly out-of-sample prediction performance (percentage R^2_{oos})

	REITs	*Holdcos*	*Banks*	*Other Fin. Insti.*	*Agriculture*
OLS	-6.89	-34.58	-9.76	-12.51	-783.31
OLS-2	0.36	0.11	-0.42	0.21	-0.40
OLS-3	0.31	0.09	-0.43	0.19	-0.42
LASSO	2.49	0.14	-0.41	0.35	-0.10
ENet	3.37	0.30	-0.29	1.10	0.75
PCR	0.28	-0.16	-0.48	0.20	-0.42
RF	2.71	-2.71	-2.34	-88.06	-2.87
GBRT	2.70	4.33	0.12	1.67	-1.60
ERT	4.52	3.95	0.40	1.10	-0.77
NN1	5.01	2.92	1.08	2.38	0.16
NN2	2.09	1.66	0.25	1.53	0.08
NN3	1.02	1.80	0.34	0.62	0.23
NN4	0.75	1.08	-0.73	0.51	-0.01
NN5	0.79	1.06	-0.34	0.81	-0.37
	Mining	*Construction*	*Other Mfr.*	*Chemicals*	*IT*
OLS	-7.58	-51.99	-3.66	-5.59	-37.34
OLS-2	-0.24	0.02	0.23	0.08	0.08
OLS-3	-0.22	-0.03	0.20	0.07	0.07
LASSO	0.11	0.05	0.54	0.15	0.30
ENet	0.63	0.52	1.33	0.42	0.87
PCR	-0.24	-0.16	0.17	0.01	0.04
RF	-0.05	-9.04	-0.01	-2.10	-3.87
GBRT	0.82	-0.41	1.34	-0.47	-0.02

ERT	0.90	0.03	1.44	0.39	0.74	
NN1	0.80	1.04	1.83	0.63	1.92	
NN2	-0.01	0.21	2.19	0.59	0.88	
NN3	-0.40	0.19	2.49	0.39	0.61	
NN4	-0.28	0.09	0.14	0.06	-0.22	
NN5	-0.06	0.05	-0.84	-0.06	-0.49	
	Transportation	*Utilities*	*Wholesale*	*Retail*	*Services*	*All Stocks*
OLS	-21.57	-14.11	-18.7	-7.87	-6.90	-2.92
OLS-2	-0.13	0.26	0.22	0.17	0.16	0.08
OLS-3	-0.07	0.23	0.20	0.09	0.13	0.06
LASSO	0.52	0.36	0.76	0.24	0.82	0.21
ENet	1.38	1.04	1.64	1.02	1.34	0.71
PCR	-0.09	0.26	0.19	0.08	0.11	0.07
RF	-10.76	-3.42	-0.89	-0.49	-0.29	0.54
GBRT	0.00	0.97	0.39	-2.24	1.23	0.25
ERT	1.03	1.50	0.86	1.03	1.32	1.03
NN1	1.32	2.20	1.96	1.34	1.12	0.28
NN2	0.72	0.85	0.75	0.39	0.76	0.27
NN3	0.34	0.55	0.58	0.51	0.71	0.32
NN4	0.18	0.51	0.36	0.19	-0.52	0.20
NN5	0.13	0.31	0.26	0.03	-0.46	0.00

Notes: This table reports monthly R^2_{oos} for the entire panel of REITs and stocks using OLS with all variables (OLS), OLS using only size and book-to-market (OLS-2), OLS using only size, book-to-market, and 12-month momentum (OLS-3), least absolute shrinkage and selection operator (LASSO), elastic net (ENet), principal component regression (PCR), random forest (RF), gradient boosted regression trees (GBRT), extremely randomized trees (ERT), and neural networks with one to five layers (NN1–NN5). All the numbers are expressed as a percentage.

Table 4: Industry characteristics

	Market capitalization ($ billion)				*Trading Volume ($ million)*			
	1990s	*2000s*	*2010s*	*Test Period*	*1990s*	*2000s*	*2010s*	*Test Period*
REIT	0.38	1.34	3.91	3.75	0.22	2.24	6.04	6.04
Investment Hold. Co.	0.22	0.49	0.68	0.75	0.15	1.41	1.13	1.66
Banking	0.66	1.82	4.12	3.83	0.36	1.52	3.30	3.27
Other Financials	1.12	4.40	6.32	6.44	0.69	4.81	7.58	8.08
Agriculture	0.43	1.67	3.69	3.48	0.27	2.67	5.40	5.45
Mining	0.56	2.71	4.22	4.11	0.46	5.77	8.64	8.89
Construction	0.20	1.04	1.72	1.86	0.20	3.68	4.84	5.50
Other Manufacturing	0.86	2.24	4.86	4.72	0.59	3.13	7.32	7.15
Chemicals	2.25	4.43	6.21	6.03	1.05	4.34	7.09	7.12
IT	0.73	2.11	5.23	5.44	0.86	4.28	6.52	6.98
Transportation	0.82	1.85	4.55	4.31	0.70	3.27	6.99	7.48
Utilities	2.08	4.14	7.90	7.49	1.07	4.46	8.90	8.61
Wholesale	0.32	1.10	2.58	2.40	0.30	1.81	4.08	3.87
Retail	0.96	2.95	7.57	7.71	0.74	4.45	10.96	10.93
Services	0.63	1.68	4.95	5.19	0.62	2.64	6.70	6.89
All Stocks	0.91	2.35	4.57	4.45	0.60	3.15	6.00	6.11

Notes: This table reports the average market capitalization ($ billions) and the average trading volume ($ millions) of REITs and stocks across time. Test period refers to the years 2006 through 2021 when out-of-sample predictability tests are performed in Section 4.2.

Table 5: Monthly out-of-sample REIT-level prediction performance by size (percentage R^2_{oos})

	All	Small	Large	Large on Small	Small on Large
OLS-2	0.36	0.17	0.69	0.22	0.25
OLS-3	0.31	0.16	0.60	0.23	0.25
LASSO	2.49	1.01	4.61	1.68	3.06
ENet	3.37	1.58	6.33	2.28	4.36
PCR	0.28	0.04	0.60	0.19	0.04
RF	2.71	0.32	4.40	2.39	1.22
GBRT	2.70	1.51	4.14	1.53	3.13
ERT	4.52	1.78	7.06	3.32	5.61
NN1	5.01	1.31	5.12	2.08	2.59
NN2	2.09	0.56	1.48	0.57	1.02
NN3	1.02	0.46	1.00	0.39	1.04
NN4	0.75	0.50	0.76	0.27	1.02
NN5	0.79	0.47	0.72	0.26	1.04

Notes: This table reports monthly $R^2_{oos}$for the entire panel of REITs, sorted by market capitalization. "Small" refers to REITs whose market capitalization are in the bottom 30th percentile, while "Large" refers to REITs whose market capitalization are in the top 30th percentile. "Large on Small" displays the R^2_{oos} of small REITs using models that are trained on large REITs. "Small on Large" displays the R^2_{oos} of large REITs using models trained on small REITs. All the numbers are expressed as a percentage.

Table 6: Monthly out-of-sample REIT-level prediction performance by property type (percentage R^2_{oos})

	Retail	Residential	Office	Healthcare	Industrial	Hotel	Diversified	Others
OLS-2	0.64	1.58	0.24	1.45	2.35	0.08	0.43	0.09
OLS-3	0.62	1.60	0.20	1.40	2.22	0.08	0.39	0.04
LASSO	2.17	3.76	3.92	4.06	4.91	8.02	1.55	1.95
ENet	3.56	5.25	5.45	4.55	5.40	8.65	3.57	2.79
PCR	0.47	1.57	0.01	1.39	1.79	-0.08	0.35	0.04
RF	4.86	6.44	4.11	3.97	0.50	3.77	-1.37	1.53
GBRT	8.21	4.04	6.54	2.60	0.24	1.51	1.17	3.18
ERT	3.38	6.07	4.32	8.16	3.49	6.19	1.60	3.27
NN1	1.97	2.79	2.10	2.91	2.47	3.15	1.62	3.24
NN2	1.00	1.86	1.04	1.51	1.93	1.21	0.81	0.97
NN3	0.82	1.49	0.65	1.49	1.70	0.70	0.66	0.39
NN4	0.71	1.81	0.27	1.51	1.74	0.59	0.78	0.36
NN5	0.79	1.73	0.23	1.52	1.48	0.32	0.70	0.40

Notes: This table reports monthly R^2_{oos} for the entire panel of REITs by property type. All the numbers are expressed as a percentage. Property type breakdowns are obtained from the S&P Global Market Intelligence database, formerly S&P Capital IQ and SNL Financial.

Table 7: Comparison of monthly out-of-sample prediction performance between large and small predictor sets (percentage R^2_{oos})

Model	*REITs*		*Stocks*	
	Full set	Reduced set	Full set	Reduced set
ENet	3.37	0.35	0.71	0.12
ERT	4.52	2.86	1.03	0.81
NN1	5.01	5.01	0.28	0.23

Notes: This table reports monthly R^2_{oos} for the entire panel of REITs and stocks using elastic net (ENet), extremely randomized trees (ERT), and neural networks with one layer (NN1). The first and third columns report the R^2_{oos} using the full set of 863 predictors described in Section 3. The second and fourth columns report the R^2_{oos} using a reduced set of 54 predictors made up of size, book-to-market and four momentum factors, and their interactions with the eight macroeconomic variables described in Section 4.9. All the numbers are expressed as a percentage.

Table 8: Relative variable importance for macroeconomic predictors

	OLS	*LASSO*	*ENet*	*PCR*	*RF*	*ERT*	*GBRT*	*NN1*	*NN2*	*NN3*	*NN4*	*NN5*
dp	56.0	0	0.0	27.5	1.6	2.2	1.6	-0.4	1.1	0.7	0.6	0.7
ep	11.1	0	0.0	37.1	42.9	68.7	1.3	1.4	1.6	2.1	1.0	0.5
bm	4.5	0	0.0	-7.0	1.0	1.4	1.3	-0.4	2.9	3.5	-0.8	0.2
ntis	0.7	0	0.2	4.6	2.9	4.3	5.6	1.9	2.2	2.6	3.2	2.8
tbl	7.5	0	0.0	3.1	1.9	2.5	1.1	0.3	1.2	1.9	0.4	-0.1
tms	4.5	0	0.9	-2.8	1.6	1.8	2.5	4.6	2.4	0.9	3.6	1.4
dfy	12.5	0	0.3	-7.5	7.5	4.0	3.3	27.0	11.8	-0.6	15.5	13.1
svar	3.1	100	98.5	45.0	40.6	15.1	83.4	65.5	76.8	88.8	76.5	81.3

Notes: This table reports the variable importance for eight macroeconomic variables detailed in Welch and Goyal (2008). These variables are dividend-to-price ratio (dp), earnings-to-price ratio (ep), book-to-market ratio (bm), net equity expansion (ntis), Treasury-bill rate (tbl), term spread (tms), default spread (dfy), and stock variance (svar). For each model, variable importance is an average over all training samples. Variable importance within each model is normalized to the sum of one. All the numbers are expressed as a percentage.

Table 9: Monthly out-of-sample prediction performance at the portfolio level (percentage R^2_{oos})

Model	*REITs, value-weighted*	*REITs, equally-weighted*
OLS-2	1.52	1.20
OLS-3	1.36	1.20
LASSO	7.33	8.08
ENet	10.79	11.36
PCR	1.79	1.20
RF	8.32	9.38
GBRT	9.41	9.71
ERT	11.46	14.38
NN1	11.86	16.91
NN2	4.78	6.43
NN3	2.49	3.22
NN4	1.87	2.43
NN5	2.25	2.56

Notes: This table reports monthly R^2_{oos} of value-weighted and equally-weighted portfolio of REITs, constructed by aggregating bottom-up forecasts of individual REIT returns and comparing them to realized portfolio returns. All the numbers are expressed as a percentage.

Table 10: Performance of value-weighted machine learning portfolios

	Avg	SD	S.R.	t-stat	Skew.	Kurt.	Max DD	Max 1M Loss	Corr
Panel A: Long-only, value-weighted portfolio									
All REITs	0.81	5.70	0.49	6.82	-0.97	5.91	-66.63	-28.25	1.00
ENet	0.83	5.52	0.52	7.18	-1.07	5.14	-57.14	-26.46	0.93
ERT	0.98	6.16	0.55	7.63	-0.47	2.39	-56.58	-24.13	0.90
NN1	0.94	5.62	0.58	8.04	-0.47	1.82	-58.66	-21.49	0.94
Nonlinear Ensemble	1.01	5.83	0.60	8.29	-0.52	3.24	-64.22	-25.52	0.96
Panel B: Long-short, value-weighted portfolio									
OLS-2	-0.09	4.76	-0.06	-0.88	-0.70	7.89	-53.42	-28.90	0.23
ENet	-0.15	4.78	-0.11	-1.49	-5.90	59.12	-68.24	-49.26	-0.38
ERT	0.41	4.13	0.35	4.79	-0.59	9.11	-40.14	-21.42	-0.32
NN1	0.35	4.10	0.29	4.05	0.80	9.40	-34.46	-17.93	-0.26
Nonlinear Ensemble	0.38	2.67	0.50	6.91	2.42	16.87	-13.80	-8.44	-0.30

Notes: This table reports the out-of-sample performance measures for the best performing machine learning models of the value-weighted long-only and long-short portfolios based on the full sample. “Avg”: average realized monthly return (%). “Std”: the standard deviation of realized monthly returns (%). “S.R.”: annualized Sharpe ratio. “T-stat”: t-statistic of realized monthly returns. “Skew”: skewness. “Kurt”: kurtosis. “Max DD”: the portfolio maximum drawdown (%). “Max 1M Loss”: the most extreme negative realized monthly return (%). “Corr”: correlation of realized monthly returns against the All REITs benchmark returns. In Panel A, the portfolios are based on a long-only strategy of holding REITs with the highest expected returns (top 30 percent), and the benchmark portfolio is the weighted index of all REITs in the sample period. In Panel B, the portfolios are based on a long-short strategy of buying REITs with the highest expected returns (top 30 percent) and shorting REITs with the lowest expected returns (bottom 30 percent), and the benchmark is a long-short portfolio based on predicted returns from OLS-2. Nonlinear ensemble refers to a grand ensemble of all nonlinear methods in our machine learning toolkit, comprising of RF, GBRT, ERT, NN1, NN2, NN3, NN4, and NN5. All portfolios are value-weighted.

Table 11: Descriptive statistics of realized and predicted returns

	Actual	*OLS-2*	*ENet*	*ERT*	*NN1*
mean	0.85	0.81	0.80	0.74	0.43
SD	11.83	0.15	1.04	1.36	2.14
min	-90.93	0.42	-12.19	-25.32	-53.50
25%	-3.62	0.70	0.81	0.88	0.37
50%	0.80	0.81	1.03	0.94	0.68
75%	5.23	0.92	1.15	0.96	0.96
max	290.27	1.31	1.64	4.77	2.06

Notes: This table reports the descriptive statistics of actual monthly REIT-level returns, and predicted monthly returns by the OLS-2, ENet, ERT and NN1 models. All the numbers are expressed as a percentage.

Table 12: Performance of machine learning portfolios using mean-variance optimization

	Avg	STD	S.R.	t-stat	Skew.	Kurt.	Max DD	Max 1M Loss	Corr
Naive 1/*N*	0.95	6.29	0.53	7.28	-0.59	7.20	-63.94	-29.27	1.00
Panel A: Long-only, mean-variance portfolio									
Sample-based	0.69	5.01	0.48	6.63	-1.03	5.12	-54.84	-23.22	0.92
ENet	0.72	5.21	0.48	6.62	-0.85	3.99	-49.86	-22.92	0.90
ERT	1.01	5.40	0.65	8.96	-0.57	3.16	-53.19	-21.80	0.88
NN1	0.94	5.61	0.58	8.08	-0.42	4.41	-44.25	-24.20	0.86
Nonlinear Ensemble	0.98	5.57	0.61	8.49	-0.62	3.32	-47.41	-23.73	0.88
Panel B: Long-short, mean-variance portfolio									
Sample-based	0.33	6.48	0.18	2.47	-1.24	4.02	-71.67	-29.98	0.37
ENet	1.13	8.00	0.49	6.76	0.58	7.47	-52.17	-31.83	0.26
ERT	2.09	12.88	0.56	7.81	0.09	20.85	-95.33	-90.10	0.01
NN1	1.37	10.34	0.46	6.35	1.85	12.60	-51.67	-26.17	0.00
Nonlinear Ensemble	1.47	9.46	0.54	7.44	2.60	26.21	-72.79	-36.71	0.07

Notes: This table reports the out-of-sample performance measures for the best performing machine learning models using mean-variance optimization. The naive strategy involves holding a portfolio weight of 1/*N* in each of the *N* REITs. In Panel A, the mean-variance portfolios are constrained to long-only positions to allow for an apples-to-apples comparison to the naive 1/*N* portfolio. In Panel B, the mean-variance portfolios are permitted to take long-short positions. “Avg”: average realized monthly return (%). “Std”: the standard deviation of realized monthly returns (%). “S.R.”: annualized Sharpe ratio. “T-stat”: t-statistic of realized monthly returns. “Skew”: skewness. “Kurt”: kurtosis. “Max DD”: the portfolio maximum drawdown (%). “Max 1M Loss”: the most extreme negative realized monthly return (%). “Corr”: correlation of realized monthly returns against the naive 1/*N* portfolio returns. Nonlinear ensemble refers to a grand ensemble of all nonlinear methods in our machine learning toolkit, comprising of RF, GBRT, ERT, NN1, NN2, NN3, NN4, and NN5.

Table 13: Time decay of predictability

	1M	3M	6M	12M
ENet	3.37	1.20	-1.22	-4.98
ERT	4.52	3.23	2.07	0.66
NN1	5.01	0.37	-0.34	-0.63

Notes: This table reports the out-of-sample R^2 of predicted returns of top machine learning models that are trained on monthly returns, versus actual returns over the next one month (1M), next three months (3M), next six months (6M) and next twelve months (12M).

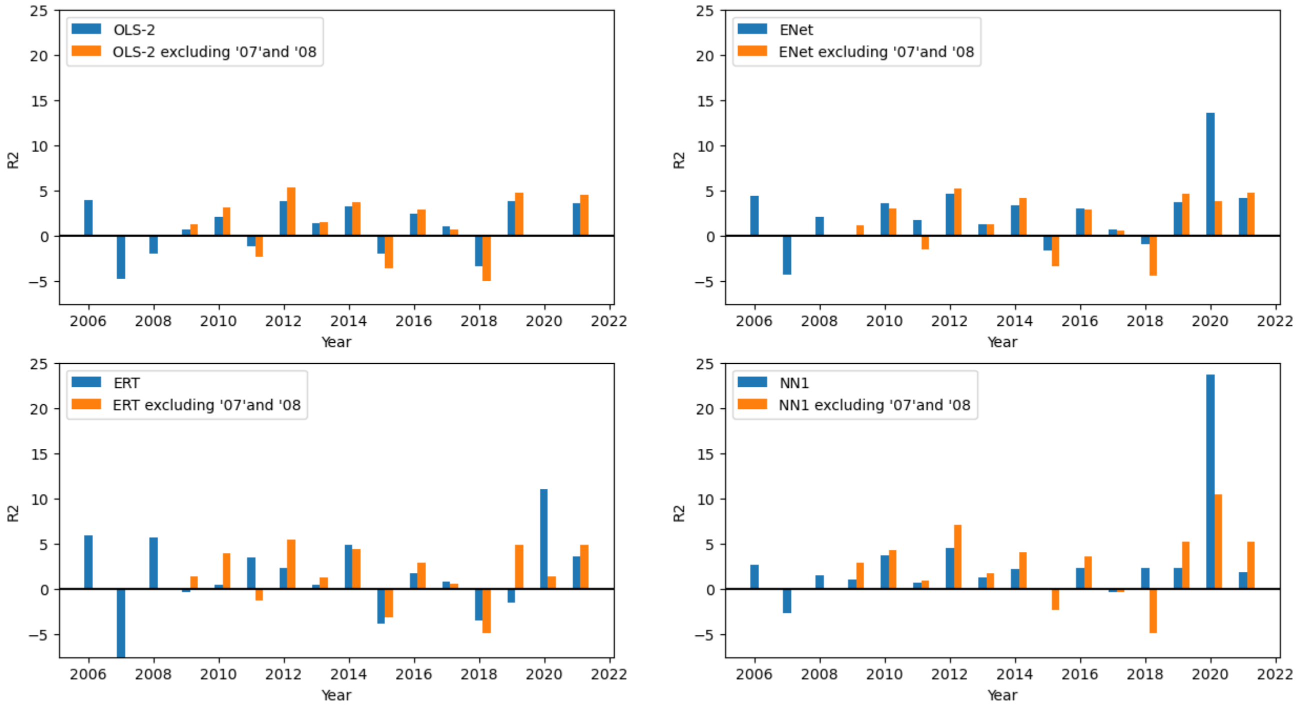
Table 14: Out-of-sample predictability with varying objective functions

	1M	3M	6M	12M
ENet	3.37	0.96	2.00	3.30
ERT	4.52	-0.72	-0.48	3.61
NN1	5.01	3.54	3.62	1.44

Notes: This table reports the out-of-sample R^2 of predicted returns versus actual returns, of top machine learning models that are trained on monthly returns (1M), quarterly returns (3M), biannual returns (6M) and annual returns (12M).

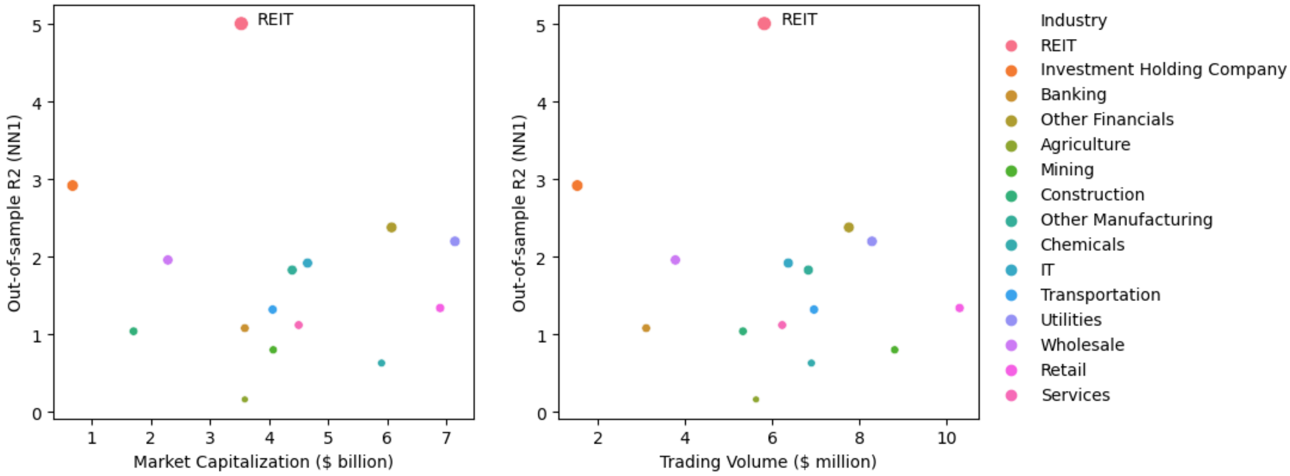
Figures

Figure 1: Out-of-sample predictive R^2, by year.



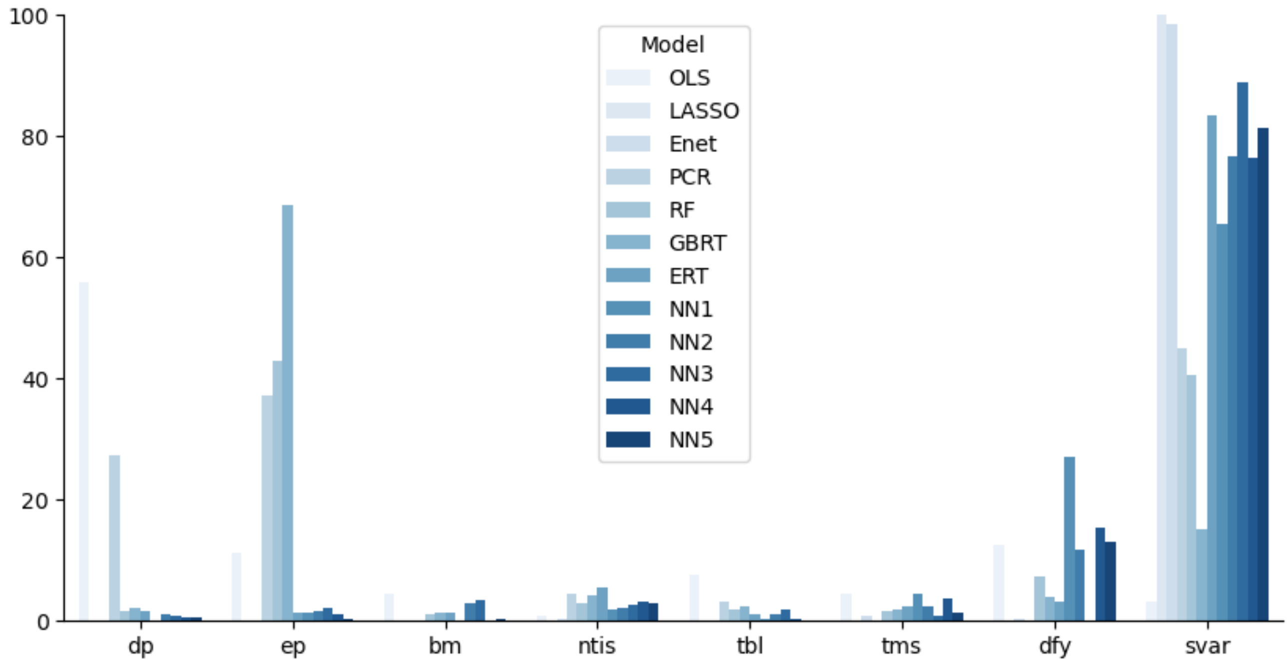
Notes: The blue bars in this figure show the monthly out-of-sample predictive R^2, averaged by calendar year, for OLS-2 and our top machine learning models (ENet, ERT and NN1) during the test period from 2006 through 2021. The orange bars display the out-of-sample predictive R^2 for each model averaged by calendar year, but after excluding the GFC years 2007 and 2008 from the training set.

Figure 2: Industry characteristics.



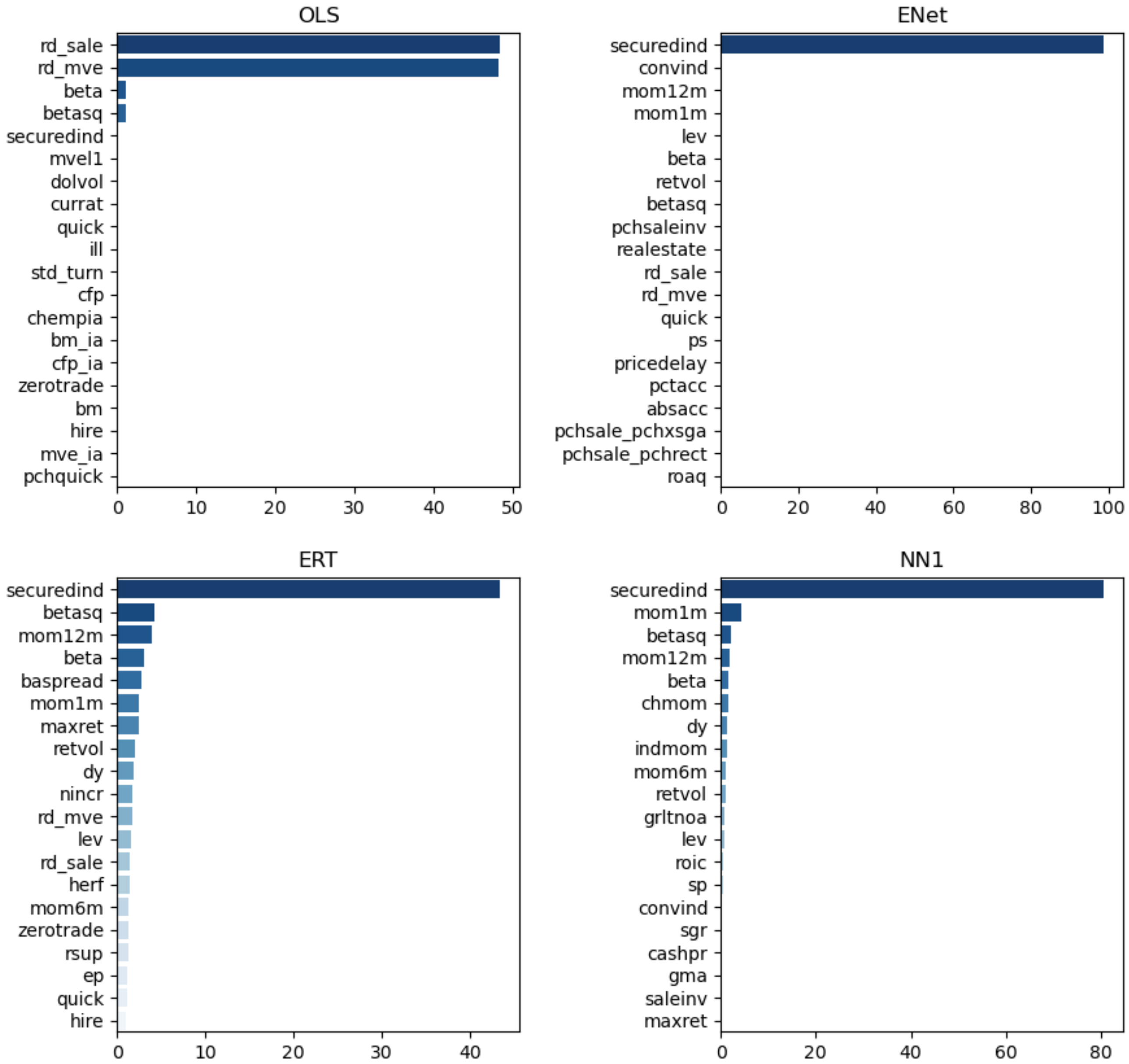
Notes: The figure on the left plots the average market capitalization of REITs and various industry sectors against the out-of-sample R^2 of the NN1 model reported in Table 3. The figure on the right plots the average trading volume of REITs and various industry sectors against the out-of-sample R^2 of the NN1 model reported in Table 3.

Figure 3: Relative variable importance for macroeconomic predictors.



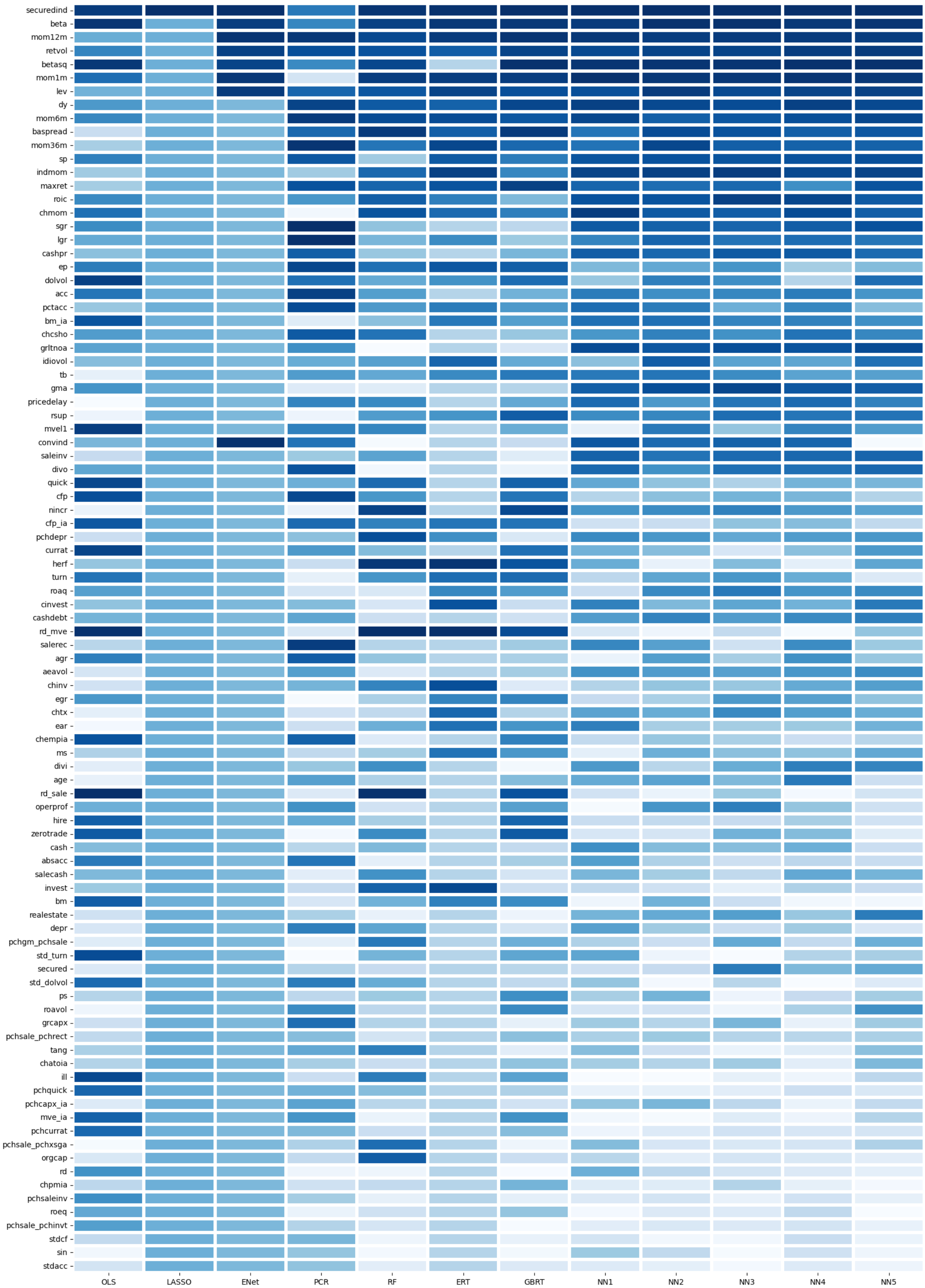
Notes: This figure provides a complementary visual comparison of the macroeconomic variable comparison across models shown in Table 8. It shows that *svar* is the most important macroeconomic variable for the regression trees and neural networks, followed by *dfy* as the second most important macroeconomic variable.

Figure 4: Relative variable importance for REIT characteristics.



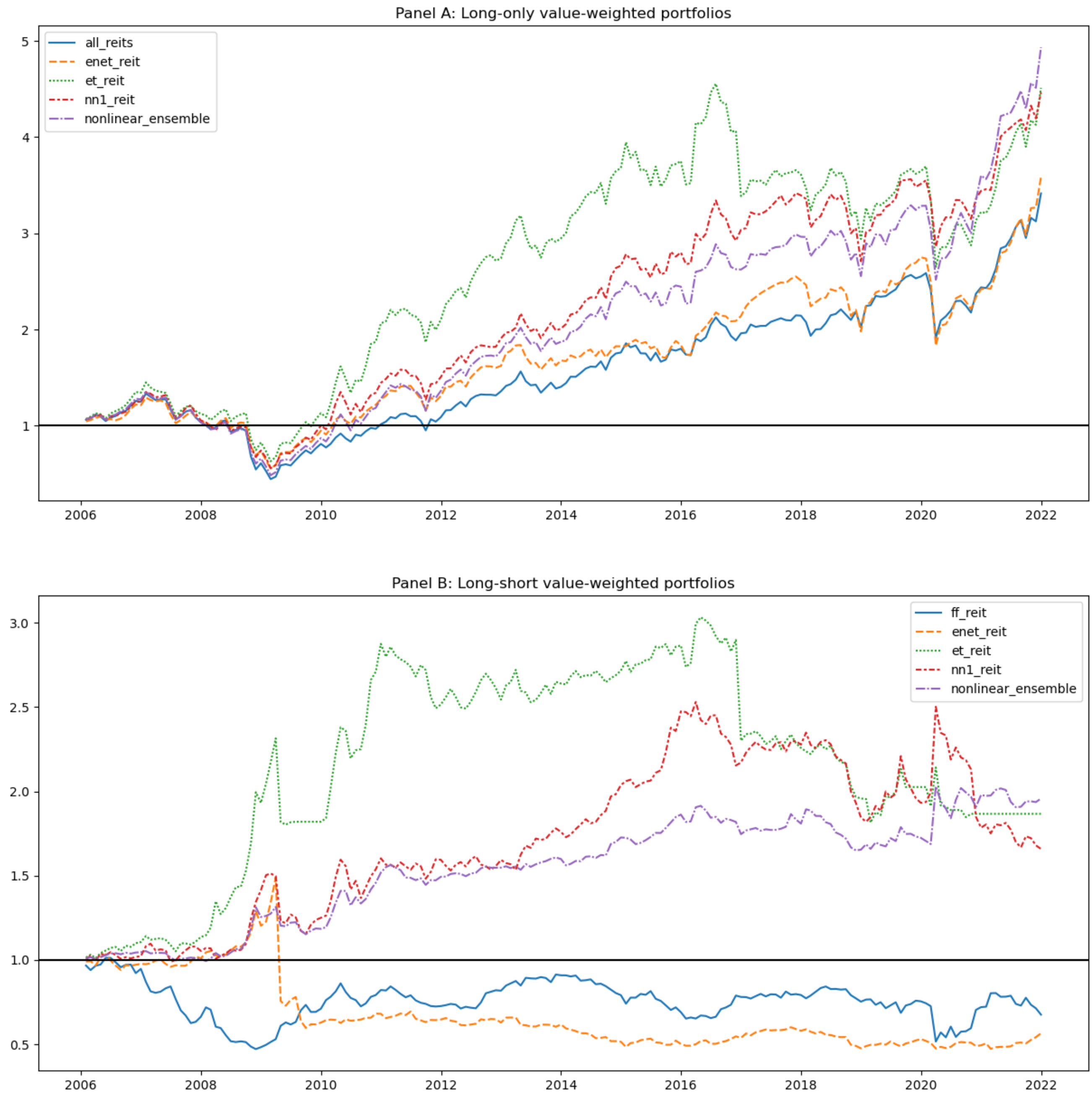
Notes: Variable importance for the top 20 most influential variables in each model. Variable importance is an average over all training samples. Variable importance within each model is normalized to sum to one. The full description of these predictors and their references are found in Appendix A. All the numbers are expressed as a percentage.

Figure 5: Relative variable importance for REIT characteristics.



Notes: This heat map shows the rankings of 94 REIT-level characteristics in terms of overall model contribution. Characteristics are ordered based on the sum of their ranks over all models, with the most influential characteristics on the top and the least influential on the bottom. Columns correspond to the individual models, and the color gradients within each column indicate the most influential (dark blue) to the least influential (white) variables. The full description of these predictors and their references is found in Appendix A.

Figure 6: Cumulative return of value-weighted machine learning portfolios.



Notes: This figure shows the cumulative returns of the best performing machine learning portfolios. In the top panel, the portfolios are based on a long-only strategy of holding REITs with the highest expected returns (top 30 percent), and the benchmark portfolio is the weighted index of all REITs in the sample period. In the bottom panel, the portfolios are based on a long-short strategy of buying REITs with the highest expected returns (top 30 percent) and shorting REITs with the lowest expected returns (bottom 30 percent), and the benchmark is a long-short portfolio based on predicted returns from OLS-2. Nonlinear_ensemble refers to a grand ensemble of all nonlinear methods in our machine learning toolkit, comprising RF, GBRT, ERT, NN1, NN2, NN3, NN4, and NN5. All portfolios are value-weighted.

APPENDICES

A Details of firm-level REIT characteristics

Table A.1: Details of firm-level characteristics

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
1	absacc	Absolute accruals	Bandyopadhyay, Huang & Wirjanto	The accrual volatility anomaly	2010, WP	Annual
2	acc	Working capital accruals	Sloan	Do stock prices fully reflect information in accruals and cash flows about future earnings?	1996, TAR	Annual
3	aeavol	Abnormal earnings announcement volume	Lerman, Livnat & Mendenhall	The high-volume return premium and post-earnings announcement drift	2008, WP	Quarterly
4	age	Number of years since first Compustat coverage	Jiang, Lee & Zhang	Information uncertainty and expected returns	2005, RAS	Annual

Note: This table lists the characteristics that we use in the empirical study. The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
5	agr	Asset growth	Cooper, Gulen & Schill	Asset growth and the cross section of asset returns	2008, JF	Annual
6	baspread	Bid-ask spread	Amihud & Mendelson	The effects of beta, bid-ask spread, residual risk, and size on stock returns	1989, JF	Monthly
7	beta	Beta	Fama & Mac-Beth	Risk, return, and equilib-rium: Empirical tests	1973, JPE	Monthly
8	betasq	Beta squared	Fama & Mac-Beth	Risk, return, and equilib-rium: Empirical tests	1973, JPE	Monthly
9	bm	Book-to-market	Rosenberg, Reid & Lanstein	Persuasive evidence of market in-efficiency	1985, JPM	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
10	bm_ia	Industry-adjusted book to market	Asness, Porter & Stevens	Predicting stock returns using industry-relative firm characteristics	2000, WP	Annual
11	cash	Cash holdings	Palazzo	Cash holdings, risk, and expected returns	2012, JFE	Quarterly
12	cashdebt	Cash flow to debt	Ou & Penman	Financial statement analysis and the prediction of stock returns	1989, JAE	Annual
13	cashpr	Cash productivity	Chandrashekar & Rao	The productivity of corporate cash holdings and the cross section of expected stock returns	2009, WP	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
14	cfp	Cash flow to price ratio	Desai, Rajgopal & Venkatachalam	Value-glamour and accruals mispricing: One anomaly or two?	2004, TAR	Annual
15	cfp_ia	Industry-adjusted cash flow to price ratio	Asness, Porter & Stevens	Predicting stock returns using industry-relative firm characteristics	2000, WP	Annual
16	chatoia	Industry-adjusted change in asset turnover	Soliman	The use of DuPont analysis by market participants	2008, TAR	Annual
17	chcsho	Change in shares outstanding	Pontiff & Woodgate	Share issuance and cross-sectional returns	2008, JF	Annual
18	chempia	Industry-adjusted change in employees	Asness, Porter& Stevens	Predicting stock returns using industry-relative firm characteristics	2000, WP	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
19	chinv	Change in inventory	Thomas & Zhang	Inventory changes and future returns	2002, RAS	Annual
20	chmom	Change in 6-month momentum	Gettleman & Marks	Acceleration strategies	2006, WP	Monthly
21	chpmia	Industry-adjusted change in profit margin	Soliman	The use of DuPont analysis by market participants	2008, TAR	Annual
22	chtx	Change in tax expense	Thomas & Zhang	Tax expense momentum	2011, JAR	Quarterly
23	cinvest	Corporate investment	Titman, Wei & Xie	Capital investments and stock returns	2004, JFQA	Quarterly
24	convind	Convertible debt indicator	Valta	Strategic default, debt structure, and stock returns	2016, JFQA	Annual

Note: This table lists the characteristics that we use in the empirical study. The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
25	currat	Current ratio	Ou & Penman	Financial statement analysis and the prediction of stock returns	1989, JAE	Annual
26	depr	Depreciation / PP&E	Holthausen & Larcker	The prediction of stock returns using financial statement information	1992, JAE	Annual
27	divi	Dividend initiation	Michaely, Thaler & Womack	Separating winners from losers among low book-to-market stocks using financial statement analysis	1995, JF	Annual
28	divo	Dividend omission	Michaely, Thaler & Womack	Separating winners from losers among low book-to-market stocks using financial statement analysis	1995, JF	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
29	dolvol	Dollar trading volume	Chordia, Subrahmanyam & Anshuman	Market liquidity and trading activity	2001, JFE	Monthly
30	dy	Dividend to price	Litzenberger & Ramaswamy	The effects of dividends on common stock prices: Tax effects or information effects?	1982, JF	Annual
31	ear	Earnings announcement return	Kishore, Brandt, Santa-Clara & Venkatachalam	Earnings announcements are full of surprises	2008, WP	Quarterly
32	egr	Growth in common shareholder equity	Richardson, Sloan, Soliman & Tuna	Accrual reliability, earnings persistence and stock prices	2005, JAE	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
33	ep	Earnings to price	Basu	Investment performance of common stocks in relation to their price-earnings ratios: A test of market efficiency	1977, JF	Annual
34	gma	Gross profitability	Novy-Marx	A taxonomy of anomalies and their trading costs	2013, JFE	Annual
35	grCAPX	Growth in capital expenditures	Anderson & Garcia-Feijoo	Empirical evidence on capital investment, growth options, and security returns	2006, JF	Annual

Note: This table lists the characteristics that we use in the empirical study. The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
36	grltnoa	Growth in long term net operating assets	Fairfield, Whisenant & Yohn	Accrued earnings and growth: Implications for future earnings performance and market mispricing	2003, TAR	Annual
37	herf	Industry sales concentration	Hou & Robinson	Industry concentration and average stock returns	2006, JF	Annual
38	hire	Employee growth rate	Bazdresch, Belo & Lin	Labor hiring, investment, and stock return predictability in the cross section	2014, JPE	Annual
39	idiovol	Idiosyncratic return volatility	Ali, Hwang & Trombley	Arbitrage risk and the book-to-market anomaly	2003, JFE	Monthly

Note: This table lists the characteristics that we use in the empirical study. The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
40	ill	Illiquidity	Amihud	Illiquidity and stock returns: cross-section and time-series effects	2002, JFM	Monthly
41	indmom	Industry momentum	Moskowitz & Grinblatt	Do industries explain momentum	1999, JF	Monthly
42	invest	Capital expenditures and inventory	Chen & Zhang	A better three-factor model that explains more anomalies	2010, JF	Annual
43	lev	Leverage	Bhandari	Debt/equity ratio and expected stock returns: Empirical evidence	1988, JF	Annual
44	lgr	Growth in long-term debt	Richardson, Sloan, Soliman & Tuna	Accrual reliability, earnings persistence and stock prices	2005, JAE	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
45	maxret	Maximum daily return	Bali, Cakici & Whitelaw	Maxing out: Stocks as lotteries and the cross section of expected returns	2011, JFE	Monthly
46	mom12m	12-month momentum	Jegadeesh	Evidence of predictable behavior of security returns	1990, JF	Monthly
47	mom1m	1-month momentum	Jegadeesh & Titman	Returns to buying winners and selling losers: Implications for stock market efficiency	1993, JF	Monthly
48	mom36m	36-month momentum	Jegadeesh & Titman	Returns to buying winners and selling losers: Implications for stock market efficiency	1993, JF	Monthly

Note: This table lists the characteristics that we use in the empirical study. The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
49	mom6m	6-month momentum	Jegadeesh & Titman	Returns to buying winners and selling losers: Implications for stock market efficiency	1993, JF	Monthly
50	ms	Financial statement score	Mohanram	Separating winners from losers among low book-to-market stocks using financial statement analysis	2005, RAS	Quarterly
51	mvel1	Size	Banz	The relationship between return and market value of common stocks	1981, JFE	Monthly
52	mve_ia	Industry-adjusted size	Asness, Porter& Stevens	Predicting stock returns using industry-relative firm characteristics	2000, WP	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
53	nincr	Number of earnings increases	Barth, Elliott & Finn	Market rewards associated with patterns of increasing earnings	1999, JAR	Quarterly
54	operprof	Operating profitability	Fama & French	A five-factor asset pricing model	2015, JFE	Annual
55	orgcap	Organizational capital	Eisfeldt & Papanikolaou	Organization capital and the cross section of expected returns	2013, JF	Annual
56	pchcapx_ia	Industry adjusted % change in capital expenditures	Abarbanell & Bushee	Abnormal returns to a fundamental analysis strategy	1998, TAR	Annual

Note: This table lists the characteristics that we use in the empirical study. The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
57	pchcurrat	% change in current ratio	Ou & Penman	Financial statement analysis and the prediction of stock returns	1989, JAE	Annual
58	pchdepr	% change in depreciation	Holthausen & Larcker	The prediction of stock returns using financial statement information	1992, JAE	Annual
59	pchgm_pch-sale	% change in gross margin - % change in sales	Abarbanell & Bushee	Abnormal returns to a fundamental analysis strategy	1998, TAR	Annual
60	pchquick	% change in quick ratio	Ou & Penman	Financial statement analysis and the prediction of stock returns	1989, JAE	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
61	pchsale_ pch- invt	% change in sales - % change in inventory	Abarbanell & Bushee	Abnormal returns to a fundamental analysis strategy	1998, TAR	Annual
62	pchsale_ pchrect	% change in sales - % change in A/R	Abarbanell & Bushee	Abnormal returns to a fundamental analysis strategy	1998, TAR	Annual
63	pchsale_ pchxsga	% change in sales - % change in SG&A	Abarbanel & Bushee	Abnormal returns to a fundamental analysis strategy	1998, TAR	Annual
64	pchsaleinv	% change sales-to- inventory	Ou & Penman	Financial statement analysis and the prediction of stock returns	1989, JAE	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
65	pctacc	Percent accruals	Hafzalla, Lund-holm & Van Winkle	Percent accruals	2011, TAR	Annual
66	pricedelay	Price delay	Hou & Moskowitz	Market frictions, price delay, and the cross section of expected returns	2005, RFS	Monthly
67	ps	Financial state-ments score	Piotroski	Value investing: The use of historical financial statement informatio n to separate winners from losers	2000, JAR	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
68	quick	Quick ratio	Ou & Penman	Financial statement analysis and the prediction of stock returns	1989, JAE	Annual
69	rd	R&D increase	Eberhart, Maxwell & Siddique	An examination of long-term abnormal stock returns and operating performance following R&D increases	2004, JF	Annual
70	rd_mve	R&D to market capitalization	Guo, Lev & Shi	Explaining the short- and long-term IPO anomalies in the US by R&D	2006, JBFA	Annual

Note: This table lists the characteristics that we use in the empirical study. The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
71	rd_sale	R&D to sales	Guo, Lev & Shi	Explaining the short- and long-term IPO anomalies in the US by R&D	2006, JBFA	Annual
72	realestate	Real estate holdings	Tuzel	Corporate real estate holdings and the cross section of stock returns	2010, RFS	Annual
73	retvol	Return volatility	Ang, Hodrick, Xing & Zhang	The cross section of volatility and expected returns	2006, JF	Monthly
74	roaq	Return on assets	Balakrishnan, Bartov & Faurel	Post loss/profit announcement drift	2010, JAE	Quarterly

Note: This table lists the characteristics that we use in the empirical study. The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
75	roavol	Earnings volatil -ity	Francis, LaFond, Olsson & Schip- per	Costs of equity and earnings attributes	2004, TAR	Quarterly
76	roeq	Return on equity	Hou, Xue & Zhang	Digesting anomalies: An investment approach	2015, RFS	Quarterly
77	roic	Return on invested capital	Brown & Rowe	The pro- ductivity premium in equity returns	2007, WP	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
78	rsup	Revenue surprise	Kama	On the market reaction to revenue and earnings surprises	2009, JBFA	Quarterly
79	salecash	Sales to cash	Ou & Penman	Financial statement analysis and the prediction of stock returns	1989, JAE	Annual
80	saleinv	Sales to inventory	Ou & Penman	Financial statement analysis and the prediction of stock returns	1989, JAE	Annual
81	salerec	Sales to receivables	Ou & Penman	Financial statement analysis and the prediction of stock returns	1989, JAE	Annual

Note: This table lists the characteristics that we use in the empirical study. The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
82	secured	Secured debt	Valta	Strategic default, debt structure, and stock returns	2016, JFQA	Annual
83	securedind	Secured debt indicator	Valta	Strategic default, debt structure, and stock returns	2016, JFQA	Annual
84	sgr	Sales growth	Lakonishok, Shleifer & Vishny	Contrarian investment, extrapolation, and risk	1994, JF	Annual
85	sin	Sin stocks	Hong & Kacperczyk	The price of sin: The effects of social norms on markets	2009, JFE	Annual

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
86	sp	Sales to price	Barbee, Mukherj, & Raines	Do the sales-price and debt-equity ratios explain stock returns better than the book-to-market value of equity ratio and firm size?	1996, FAJ	Annual
87	std_dolvol	Volatility of liquidity (dollar trading volume)	Chordia, Subrahmanyam & Anshuman	Trading activity and expected stock returns	2001, JFE	Monthly
88	std_turn	Volatility of liquidity (share turnover)	Chordia, Subrahmanyam & Anshuman	Trading activity and expected stock returns	2001, JFE	Monthly
89	stdacc	Accrual volatility	Bandyopadhyay, Huang & Wirjanto	The accrual volatility anomaly	2010, WP	Quarterly

Note: This table lists the characteristics that we use in the empirical study. The data are collected in Green et al. (2017) and Gu et al. (2020).

Table A.1: Details of firm-level characteristics (continued)

No.	Acronym	Characteristic	Paper's author(s)	Paper's title	Year, Journal	Frequency
90	stdcf	Cash flow volatility	Huang	The cross section of cash flow volatility and expected stock returns	2009, JEF	Quarterly
91	tang	Debt capacity/ firm tangibility	Almeida & Campello	Financial constraints, asset intangibility, and corporate investment	2007, RFS	Annual
92	tb	Tax income to book income	Lev & Nissim	Taxable income, future earnings, and equity values	2004, TAR	Annual
93	turn	Share turnover	Datar, Naik & Radcliffe	Liquidity and stock returns: An alternative test	1998, JFM	Monthly
94	zerotrade	Zero trading days	Liu	A liquidity-augmented capital asset pricing model	2006, JFE	Monthly

Note: This table lists the characteristics that we use in the empirical study.
The data are collected in Green et al. (2017) and Gu et al. (2020).

B Details on macroeconomic variables

Table B.1: Details of macroeconomic variables

No.	Acronym	Macro Variable	Definition	Paper's author(s)	Paper's title	Year, Journal
1	dp	Dividend price ratio	The difference between the log of dividends and the log of prices. Dividends are 12-month moving sums of dividends paid on the S&P 500 index.	Campell and Shiller	The dividend-price ratio and expectations of future dividends and discount factors	1988, RFS
2	ep	Earnings price ratio	The difference between the log of earnings and the log of prices. Earnings are 12-month moving sums of earnings on the S&P 500 index.	Campbell and Shiller	Stock prices, earnings, and expected dividends	1988, JF

Note: This table lists the macroeconomic variables that we use in the empirical study. The data are collected in Welch and Goyal (2008).

Table B.1: Details of macroeconomic variables (continued)

No.	Acronym	Macro Variable	Definition	Paper's author(s)	Paper's title	Year, Journal
3	bm	Book-to-market ratio	The ratio of book value to market value for the Dow Jones Industrial Average index	Kothari and Shanken	Book-to-market, dividend yield, and expected market returns: a time-series analysis	1997, JFE
4	ntis	Net equity expansion	The ratio of 12-month moving sums of net issues by NYSE listed stocks divided by the total end-of-year market capitalization of NYSE stocks	Boudoukh, Michaely, Richardson, and Roberts	On the importance of measuring payout yield: implications for empirical asset pricing	2007, JF
5	tbl	Treasury bill	The 3-month treasury bill rate	Campbell	Stock returns and the term structure	1987, JFE

Note: This table lists the macroeconomic variables that we use in the empirical study. The data are collected in Welch and Goyal (2008).

Table B.1: Details of macroeconomic variables (continued)

No.	Acronym	Macro Variable	Definition	Paper's author(s)	Paper's title	Year, Journal
6	tms	Term spread	The difference between the long term yield on government bonds and the Treasury bill	Campbell	Stock returns and the term structure	1987, JFE
7	dfy	Default yield spread	The difference between BAA and AAA-rated corporate bond yields	Fama and French	Business conditions and expected returns on stocks and bonds	1989, JFE
8	svar	Stock variance	The sum of squared daily returns on the S&P 500 index	Guo	On the out-of-sample predictability of stock market returns	2006, JB

Note: This table lists the macroeconomic variables that we use in the empirical study. The data are collected in Welch and Goyal (2008).

C Details on machine learning methodology

C.1 LASSO

The Least Absolute Shrinkage and Selection Operator (LASSO) methodology is a regularization technique used in machine learning for linear regression models, introduced by Tibshirani (1996) who coined the term. The goal of LASSO is to prevent overfitting, which occurs when a model is too complex and fits the training data too well but performs poorly on new data.

LASSO works by adding a penalty term to the cost function, which is a function that measures the error between the predicted values and the actual values. This penalty term is the sum of the absolute values of the coefficients of the features in the model. By adding this penalty term, LASSO encourages the coefficient estimates of less important features to be shrunk towards zero. This has the effect of removing these features from the model, which simplifies it and makes it more interpretable.

The amount of shrinkage is controlled by a tuning hyperparameter. LASSO is particularly useful when dealing with high-dimensional datasets with many features, where it can help identify the most important features for prediction. Contrary to other regularization techniques, such as Ridge regression, LASSO can lead to sparse models where many of the coefficients are exactly zero. This can be useful for feature selection, as it can identify the most important features in the dataset.

C.2 ENet

ENet, or Elastic Net, is a regularization technique used in machine learning for linear regression models, introduced by Zhou and Hastie (2005). It is a combination of LASSO and Ridge (Hoerl & Kennard, 1970) regression, which aims to address the limitations of each method. Like LASSO, ENet can lead to sparse models by shrinking less important coefficients towards zero, but it also includes a Ridge penalty term that prevents overfitting by adding a bias towards small coefficient values.

The ENet methodology adds two penalty terms to the cost function: one each for the L1 and L2 norms of the coefficients. The relative importance of the two terms is controlled by a hyperparameter alpha. When alpha is set to 1, ENet is equivalent to LASSO, and when alpha is set to 0, it is equivalent to Ridge regression. By tuning the alpha parameter, the model can strike a balance between LASSO and Ridge regression.

ENet is particularly useful when dealing with high-dimensional datasets where there are many features, some of which may be correlated. In such cases, LASSO may select only one feature from a group of correlated features, while Ridge regression may include all of them. ENet can identify the most important features while still accounting for correlations between features.

ENet is a flexible regularization technique that can lead to more accurate and interpretable models, particularly in situations where LASSO or Ridge regression alone may not be sufficient.

C.3 PCR

Principal Component Regression (PCR) is a technique used in machine learning for regression analysis. It is a form of dimensionality reduction that involves transforming the original features of the dataset into a smaller set of principal components. The principal components are then used as predictors in a regression model, instead of the original features.

PCR works by identifying the linear combinations of the original features that explain the most variation in the dataset. These linear combinations are called principal components, and they are orthogonal to each other. The first principal component explains the most variation in the data, followed by the second principal component, and so on. The number of principal components is usually chosen based on the amount of variation they explain and the desired level of model complexity.

Once the principal components are identified, they can be used as predictors in a regression model. This approach can help address the issue of multicollinearity, which occurs when the original features are highly correlated with each other. By transforming the features into principal components, PCR can reduce the number of predictors in the model while retaining the most important information about the original features.

PCR can be particularly useful when dealing with high-dimensional datasets with many features that may be redundant or correlated. By reducing the number of predictors in the model, PCR can improve the interpretability and stability of the model, and reduce the risk of overfitting.

C.4 RF

Random Forest (RF) regression is a machine learning algorithm that is used for predicting continuous values (Breiman, 2001). It is based on the same principle as Random Forest for classification, but instead of predicting a categorical outcome, it predicts a continuous value.

To create a RF regression model, the algorithm constructs multiple decision trees, where each tree is trained on a random subset of the features and a random subset of the training samples. During the training process, each decision tree makes a prediction based on a subset of the features and a subset of the training data. The final prediction is then made by aggregating the predictions of all the decision trees in the forest. In regression problems, the prediction is usually the mean of the predictions.

RF regression is particularly useful when dealing with datasets with many features and large numbers of training examples. It can handle non-linear relationships between the features and the target variable and can automatically detect interactions between the features. Additionally, RF can handle missing values and can provide an estimate of the importance of each feature in the prediction.

C.5 GBRT

Gradient Boosted Regression Trees (GBRT) is a machine learning algorithm that is used for regression problems (Buhlmann & Hothorn, 2007; Friedman, 2001). It is based on the same principle as gradient boosting, where multiple weak learners (decision trees) are combined to create a strong learner.

To train a GBRT model, the algorithm first creates a decision tree based on the training data. The errors between the predictions and the true values are then calculated, and the algorithm creates another decision tree to predict the residual errors. This process is repeated multiple times, with each new tree predicting the residual errors of the previous tree. The final prediction is then made by aggregating the predictions of all the trees.

GBRT is particularly useful when dealing with high-dimensional datasets with many features, where other models may struggle to find meaningful relationships between the features and the target variable. It is also effective at handling non-linear relationships between the features and the target variable. Like Random Forest, it can handle missing data and outliers.

C.6 ERT

Extremely Randomized Trees (ERT) regression was proposed by Geurts et al. (2006). ERT regression and RF regression are both ensemble learning algorithms used for regression problems, but the main difference between ERT and RF is the way they construct the decision trees. In RF, each decision tree is constructed with a random subset of the features and a random subset of the training samples. The best split point is then chosen from the subset of features at each node based on a splitting criterion such as information gain or Gini impurity. Conversely, ERT constructs each decision tree using random splits on random subsets of both the features and training samples.

Another difference is the level of randomness in the model. In RF, each tree is grown independently and then combined to make the final prediction. In ERT, the level of randomness is increased by using only a subset of the training samples to select the split point at each node, leading to a greater degree of diversity among the trees.

ERT has been shown to have several advantages over RF in terms of both predictive accuracy and computation time. ERT can be less prone than RF to overfitting and can handle high-dimensional data more efficiently, making it a good choice for certain types of regression problems.

C.7 NN1–NN5

Arguably the most powerful modeling (and most computationally intensive) device in machine learning (Cybenko, 1989; Hornik et al., 1989), neural network (NN) is the currently preferred approach for solving complex machine learning problems, such as computer vision and natural language processing. It is based on the structure and function of the human brain, with interconnected nodes that process information and make predictions. The neural network consists of multiple layers of nodes, where each node receives input signals from the previous layer and applies a mathematical function to the signals to produce an output. The output of the final layer represents the predicted value for the output variable.

During training, the neural network adjusts the weights of the connections between the nodes to minimize the difference between the predicted output and the actual output. This is achieved through an optimization algorithm, such as gradient descent, which adjusts the weights in the direction of the steepest descent of the cost function.

Neural network regression is particularly useful when dealing with complex, high-dimensional datasets with nonlinear relationships between the features and the target variable. It can automatically learn complex patterns and relationships between the features and the target variable, which can be difficult or impossible to discover using traditional regression models. Additionally, neural networks can handle missing data and can be trained on large datasets.

However, neural networks can be computationally expensive to train and require a large amount of data to avoid overfitting. Overfitting occurs when the neural network fits the training data too closely and fails to generalize to new, unseen data. It requires careful tuning of hyperparameters and regularization techniques to achieve optimal performance.

In our study, we employ neural networks with up to five hidden layers. Each layer consists of a certain number of neurons, which are built with the commonly used rectified linear unit (ReLU). Our shallowest neural network has a single hidden layer of 32 neurons, which we denote NN1. Next, NN2 has two hidden layers with 32 and 16 neurons, respectively; NN3 has three hidden layers with 32, 16, and 8 neurons, respectively; NN4 has four hidden layers with 32, 16, 8, and 4 neurons, respectively; and NN5 has five hidden layers with 32, 16, 8, 4, and 2 neurons, respectively. We choose the number of neurons in each layer according to the geometric pyramid rule (see Masters, 1993). We adopt the Adam optimization algorithm (Kingma and Ba, 2014), early stopping, batch normalization (Ioffe and Szegedy, 2015), ensembles, and dropout (Srivastava et al., 2014) when training our models.

D Default hyperparameters for machine learning methods

We do not require a validation sample as we do not perform any hyperparameter optimization, following Elkind et al. (2022). We employ Scikit-Learn[9], an open-source machine learning library for Python built on top of SciPy, for our linear machine methods, random forest regression and extremely randomized trees regression. Our neural networks are trained using TensorFlow[10] developed by the Google Brain team, and the Keras[11] wrapper, an open-source software library that provides a Python interface for artificial neural networks. The Keras wrapper provides two distinct approaches to constructing neural networks, i.e. a sequential API and a functional API. The sequential API constructs simple network structures that do not require merged layers, while the functional API is used to build those networks that required sophisticated merged layers. Our study uses sequential API. Keras also implements a range of regularization methods described in Appendix C, such as early-stopping, dropout, batch normalization and L1/L2 penalties.

Default hyperparameters of these models are used where possible. This forms the lowest bound of performance for our machine learning models. Machine learning training is executed on an Apple M1 Ultra chip with a 20-core CPU, a 48-core GPU and 128 GB unified memory.[12]

Table D.1: Default hyperparameters for machine learning methods

No.	Machine Learning Model	Default Hyperparameters
1	Least absolute shrinkage and selection operator (LASSO)	alpha=1.0 max_iter=1000 tol=0.0001 random_state=42

9 scikit-learn v1.2.1, https://scikit-learn.org/stable/

10 tensorflow-macos v2.9.0, https://www.tensorflow.org/

11 keras v2.9.0, https://keras.io/

12 While these CPU, GPU and memory specifications are extremely powerful for a personal computer (Apple claims the M1 Ultra is the most powerful chip ever in a personal computer, as of 1 April 2023), regression trees and neural networks do stretch the computer to its limit, even without attempting hyperparameter tuning. Equipped with a more powerful GPU such as the Nvidia Tesla K80 with thousands of cores, hyperparameter tuning can take place and we would expect better performance results for both REIT and stock predictions, but we do not expect a qualitative difference in our conclusion that REITs are more predictable than stocks.

Table D.1: Default hyperparameters for machine learning methods (continued)

No.	Machine Learning Model	Default Hyperparameters
2	Elastic net (ENet)	alpha=1.0 l1_ratio=0.5 max_iter=2000 tol=0.0001 random_state=42
3	Principle compo- nent regression (PCR)	n_components=5 tol=0.0 random_state=42
4	Random forest (RF)	n_trees=300 max_depth=2 min_samples_split=2 min_samples_leaf=1 min_weight_fraction_leaf=0.0 max_features=p/3 (recommended default value formula, per Hastie et al., 2009) max_leaf_nodes=None min_impurity_decrease=0.0 ccp_alpha=0.0 max_samples=None random_state=42
5	Gradien boosted regression trees (GBRT)	learning_rate=0.1 n_trees=100 max_leaf_nodes=31 max_depth=1 min_sample_leaf=20 l2_regulartization=0 max_bins=255

Table D.1: Default hyperparameters for machine learning methods (continued)

No.	Machine Learning Model	Default Hyperparameters
6	Extremely randomized trees (ERT)	n_trees=300 max_depth=2 min_samples_split=2 min_samples_leaf=1 min_weight_fraction_leaf=0.0 max_features=p/3 (recommended default value formula, per Hastie et al., 2009) max_leaf_nodes=None min_impurity_decrease=0.0 ccp_alpha=0.0 max_samples=None random_state=42
7	Neural network (NN)	activation=relu penalty_type=l1 penalty_amount=1 optimizer=Adam learning_rate=0.01 batch_normalization=yes early_stopping=yes min_delta=0 patience=5 epochs=100 batch_size=2^{13} ensemble=5 random_state=42

E Predictability over time

Table E.1: Monthly predictive R^2_{oos}, averaged by calendar year (percentage R^2_{oos}), excluding 2007 and 2008 as training years

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-	-	-	-36.08	-2.18	-5.46	-3.09	-17.68	-32.02
OLS-2	-	-	-	1.25	3.10	-2.30	5.32	1.49	3.70
OLS-3	-	-	-	0.89	3.17	-2.28	5.01	1.80	3.59
LASSO	-	-	-	1.18	2.98	-1.90	5.07	1.22	4.36
ENet	-	-	-	1.15	3.01	-1.47	5.23	1.28	4.22
PCR	-	-	-	1.27	3.41	-2.67	5.46	1.64	3.72
RF	-	-	-	1.23	2.89	-0.37	-4.16	-2.87	4.64
GBRT	-	-	-	1.74	2.09	1.70	-103.51	1.02	4.38
ERT	-	-	-	1.37	3.93	-1.21	5.47	1.24	4.38
NN1	-	-	-	2.89	4.33	0.91	7.08	1.74	4.03
NN2	-	-	-	1.16	2.46	-0.93	4.48	1.24	3.91
NN3	-	-	-	1.13	2.90	-1.01	4.44	1.24	3.94
NN4	-	-	-	1.08	2.58	-1.15	4.46	1.24	3.92
NN5	-	-	-	1.06	2.52	-1.11	4.36	1.23	3.91

	2015	2016	2017	2018	2019	2020	2021	All Years	All ex ’20-21
OLS	-22.37	-5.46	-5.28	1.24	-15.19	-5.25	-0.52	-15.8	-19.86
OLS-2	-3.63	2.97	0.73	-5.01	4.72	-0.14	4.49	1.18	1.29
OLS-3	-3.45	2.69	0.72	-4.83	4.65	-0.22	4.43	1.06	1.15
LASSO	-3.35	2.82	0.69	-4.92	4.68	-0.16	4.75	1.18	1.27
ENet	-3.38	2.92	0.61	-4.41	4.62	3.82	4.77	2.03	1.33
PCR	-3.94	2.93	0.42	-5.78	4.54	-0.42	4.33	1.08	1.24
RF	-3.97	-76.14	0.67	-9.87	4.85	1.18	4.81	-3.41	-5.30
GBRT	-4.82	-33.64	0.76	-2.03	-3.72	6.47	2.92	-4.41	-7.96
ERT	-3.16	2.91	0.60	-4.88	4.83	1.36	4.83	1.69	1.54
NN1	-2.27	3.64	-0.39	-4.86	5.29	10.49	5.22	4.29	2.53
NN2	-1.99	2.54	1.01	-3.94	4.57	1.19	4.53	1.49	1.33
NN3	-2.36	2.95	1.02	-4.32	4.81	0.01	4.68	1.28	1.36
NN4	-2.70	2.91	0.95	-5.27	4.85	-0.21	4.66	1.14	1.23
NN5	-2.63	2.78	0.89	-4.49	4.89	-0.16	4.49	1.15	1.25

Notes: This table reports the monthly out-of-sample R^2 for the entire panel of REITs, averaged by calendar year, but excluding calendar years 2007 and 2008 from the training data.

F Monthly predictive R^2_{oos}, by industry and by year

Figure F.1: Industries split by SIC codes

2-Digit SIC (Standard Industrial Classification) Codes

A. Agriculture, Forestry, & Fishing
01 Agricultural Production – Crops
02 Agricultural Production – Livestock
07 Agricultural Services
08 Forestry
09 Fishing, Hunting, & Trapping

B. Mining
10 Metal, Mining
12 Coal Mining
13 Oil & Gas Extraction
14 Nonmetallic Minerals, Except Fuels

C. Construction
15 General Building Contractors
16 Heavy Construction, Except Building
17 Special Trade Contractors

D. Manufacturing
20 Food & Kindred Products
21 Tobacco Products
22 Textile Mill Products
23 Apparel & Other Textile Products
24 Lumber & Wood Products
25 Furniture & Fixtures
26 Paper & Allied Products
27 Printing & Publishing
28 Chemical & Allied Products
29 Petroleum & Coal Products
30 Rubber & Miscellaneous Plastics Products
31 Leather & Leather Products
32 Stone, Clay, & Glass Products
33 Primary Metal Industries
34 Fabricated Metal Products
35 Industrial Machinery & Equipment
36 Electronic & Other Electric Equipment
37 Transportation Equipment
38 Instruments & Related Products
39 Miscellaneous Manufacturing Industries

E. Transportation & Public Utilities
40 Railroad Transportation
41 Local & Interurban Passenger Transit
42 Trucking & Warehousing
43 U.S. Postal Service
44 Water Transportation
45 Transportation by Air
46 Pipelines, Except Natural Gas
47 Transportation Services
48 Communications
49 Electric, Gas, & Sanitary Services

F. Wholesale Trade
50 Wholesale Trade – Durable Goods
51 Wholesale Trade – Nondurable Goods

G. Retail Trade
52 Building Materials & Gardening Supplies
53 General Merchandise Stores
54 Food Stores
55 Automative Dealers & Service Stations
56 Apparel & Accessory Stores
57 Furniture & Homefurnishings Stores
58 Eating & Drinking Places
59 Miscellaneous Retail

H. Finance, Insurance, & Real Estate
60 Depository Institutions
61 Nondepository Institutions
62 Security & Commodity Brokers
63 Insurance Carriers
64 Insurance Agents, Brokers, & Service
65 Real Estate
67 Holding & Other Investment Offices

I. Services
70 Hotels & Other Lodging Places
72 Personal Services
73 Business Services
75 Auto Repair, Services, & Parking
76 Miscellaneous Repair Services
78 Motion Pictures
79 Amusement & Recreation Services
80 Health Services
81 Legal Services
82 Educational Services
83 Social Services
84 Museums, Botanical, Zoological Gardens
86 Membership Organizations
87 Engineering & Management Services
88 Private Households
89 Services, Not Elsewhere Classified

J. Public Administration
91 Executive, Legislative, & General
92 Justice, Public Order, & Safety
93 Finance, Taxation, & Monetary Policy
94 Administration of Human Resources
95 Environmental Quality & Housing
96 Administration of Economic Programs
97 National Security & International Affairs
98 Zoological Gardens

K. Nonclassifiable Establishments
99 Non-Classifiable Establishments

Notes: List of major industries and their respective 2-digit SIC codes from the United States Department of Labor.

Table F.1: Monthly out-of-sample prediction performance (percentage R^2_{oos})

	REITs		*Holdcos*		*Banks*		*Other Fin. Insti.*		*Agriculture*	
	All Years	*2020*	*All Years*	*2020*	*All Years*	*2020*	*All Years*	*2020*	*All Years*	*2020*
OLS	-6.89	7.82	-34.58	-15.14	-9.76	-9.49	-12.51	-13.69	-783.31	-1680.19
OLS-2	0.36	-0.04	0.11	0.38	-0.42	-0.87	0.21	0.62	-0.40	0.34
OLS-3	0.31	-0.15	0.09	0.34	-0.43	-0.96	0.19	0.56	-0.42	0.34
LASSO	2.49	11.69	0.14	0.35	-0.41	-0.86	0.35	2.31	-0.10	1.82
ENet	3.37	13.61	0.30	1.01	-0.29	-0.71	1.10	4.93	0.75	9.71
PCR	0.28	-0.23	-0.16	0.29	-0.48	-0.98	0.20	0.61	-0.42	0.34
RF	2.71	12.94	-2.71	11.59	-2.34	0.10	-88.06	4.06	-2.87	1.87
GBRT	2.70	11.09	4.33	15.31	0.12	1.14	1.67	5.08	-1.60	2.61
ERT	4.52	21.13	3.95	16.43	0.40	8.35	1.10	8.68	-0.77	10.84
NN1	5.01	23.71	2.92	13.30	1.08	9.52	2.38	8.82	0.16	3.26
NN2	2.09	9.95	1.66	5.97	0.25	6.92	1.53	6.68	0.08	1.86
NN3	1.02	3.93	1.80	6.65	0.34	9.66	0.62	2.18	0.23	0.53
NN4	0.75	2.49	1.08	3.35	-0.73	-0.39	0.51	3.21	-0.01	-0.10
NN5	0.79	2.34	1.06	4.80	-0.34	8.22	0.81	7.07	-0.37	-0.06

	Mining		*Construction*		*Other Mfr.*		*Chemicals*		*IT*	
	All Years	*2020*	*All Years*	*2020*	*All Years*	*2020*	*All Years*	*2020*	*All Years*	*2020*
OLS	-7.58	-12.03	-51.99	-144.31	-3.66	-10.03	-5.59	-6.52	-37.34	-16.15
OLS-2	-0.24	0.23	0.02	1.31	0.23	0.69	0.08	0.75	0.08	1.46
OLS-3	-0.22	0.24	-0.03	1.44	0.20	0.66	0.07	0.74	0.07	1.45
LASSO	0.11	2.81	0.05	3.88	0.54	2.32	0.15	1.06	0.30	2.63
ENet	0.63	3.17	0.52	6.40	1.33	3.18	0.42	1.65	0.87	2.88
PCR	-0.24	0.15	-0.16	0.41	0.17	0.67	0.01	0.66	0.04	1.27
RF	-0.05	3.57	-9.04	2.14	-0.01	0.85	-2.10	0.34	-3.87	0.09
GBRT	0.82	2.04	-0.41	5.95	1.34	2.81	-0.47	1.30	-0.02	2.83
ERT	0.90	6.14	0.03	9.26	1.44	3.53	0.39	1.49	0.74	2.71
NN1	0.80	5.68	1.04	4.52	1.83	3.26	0.63	2.61	1.92	2.21
NN2	-0.01	2.09	0.21	1.20	2.19	3.82	0.59	1.43	0.88	1.33
NN3	-0.40	0.69	0.19	0.75	2.49	3.65	0.39	1.08	0.61	-0.03
NN4	-0.28	0.73	0.09	0.49	0.14	0.67	0.06	0.30	-0.22	-1.63
NN5	-0.06	0.26	0.05	0.73	-0.84	-2.24	-0.06	0.23	-0.49	-1.41

	Transportation		*Utilities*		*Wholesale*		*Retail*		*Services*		*All Stocks*	
	All Years	*2020*	*All Years*	*2020*	*All Years*	*2020*	*All Years*	*2020*	*All Years*	*2020*	*All Years*	*2020*
OLS	-21.57	-34.96	-14.11	-6.98	-18.7	-31.4	-7.87	-16.3	-6.90	-11.6	-2.92	-7.25
OLS-2	-0.13	-0.10	0.26	0.22	0.22	0.66	0.17	0.57	0.16	1.03	0.08	0.59
OLS-3	-0.07	-0.24	0.23	0.16	0.20	0.63	0.09	0.48	0.13	0.96	0.06	0.55
LASSO	0.52	5.05	0.36	0.50	0.76	4.02	0.24	2.16	0.82	3.15	0.21	1.24
ENet	1.38	6.13	1.04	1.04	1.64	5.92	1.02	3.59	1.34	4.02	0.71	2.17
PCR	-0.09	0.07	0.26	0.21	0.19	0.69	0.08	0.18	0.11	0.86	0.07	0.55
RF	-10.76	4.11	-3.42	0.01	-0.89	1.02	-0.49	5.68	-0.29	1.26	0.54	0.53
GBRT	0.00	5.42	0.97	1.55	0.39	4.37	-2.24	5.88	1.23	3.21	0.25	-0.36
ERT	1.03	5.67	1.50	1.85	0.86	5.55	1.03	6.42	1.32	3.50	1.03	3.14
NN1	1.32	4.81	2.20	1.79	1.96	7.31	1.34	5.22	1.12	4.83	0.28	0.8
NN2	0.72	4.97	0.85	0.96	0.75	3.00	0.39	3.41	0.76	4.55	0.27	0.59
NN3	0.34	1.47	0.55	0.45	0.58	0.77	0.51	4.33	0.71	2.57	0.32	0.59
NN4	0.18	0.77	0.51	1.07	0.36	0.55	0.19	1.29	-0.52	3.79	0.20	0.58
NN5	0.13	0.43	0.31	0.13	0.26	0.18	0.03	-0.47	-0.46	5.57	0.00	0.60

Notes: This table reports R^2_{oos} for the entire for the entire panel of REITs and stocks using OLS with all variables (OLS), OLS using only size and book-to-market (OLS-2), OLS using only size, book-to-market, and 12-month momentum (OLS-3), least absolute shrinkage and selection operator (LASSO), elastic net (ENet), principal component regression (PCR), random forest (RF), gradient boosted regression trees (GBRT), extremely randomized trees (ERT), and neural networks with one to five layers (NN1–NN5). All the numbers are expressed as a percentage.

Table F.2: Monthly predictive R^2_{oos}, by year (Investment Hold Cos)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-16.73	-330.93	-59.4	-13.83	-0.17	-8.34	-1.25	-76.63	-16.52
OLS-2	2.77	-3.14	-2.32	0.82	1.71	0.24	2.68	0.46	1.14
OLS-3	2.82	-2.64	-2.12	0.38	1.86	-0.29	2.53	1.15	0.96
LASSO	3.17	-3.05	-2.27	0.78	1.85	0.26	2.66	0.52	1.37
ENet	3.17	-3.05	-2.27	0.66	1.97	1.09	2.65	0.40	1.44
PCR	3.05	-3.00	-3.25	0.91	2.40	-1.06	2.28	0.73	0.81
RF	3.88	-2.97	-23.24	-3.41	2.20	0.13	3.06	1.02	1.38
GBRT	2.56	-2.70	3.12	3.59	3.09	0.68	3.47	3.43	0.70
ERT	3.35	-2.23	4.30	-1.09	2.62	1.53	3.18	0.45	1.49
NN1	2.44	-1.73	-1.52	2.47	2.58	0.19	3.35	0.66	0.79
NN2	2.42	-1.74	-1.49	2.47	2.53	0.19	3.37	0.51	0.53
NN3	2.46	-1.75	-1.54	2.42	2.63	0.19	3.35	0.51	0.59
NN4	2.45	-1.74	-1.56	2.45	2.57	0.19	3.36	0.51	0.68
NN5	2.47	-1.76	-1.50	2.46	2.52	0.19	3.37	0.51	0.11

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-5.09	-13.36	-1.11	-6.16	-3.72	-15.14	-3.47	-34.58
OLS-2	-1.49	1.46	3.22	-3.04	2.53	0.38	2.72	0.11
OLS-3	-0.88	0.83	3.75	-3.36	2.60	0.34	2.62	0.09
LASSO	-1.54	1.55	3.19	-2.99	2.58	0.35	2.76	0.14
ENet	-1.27	1.62	3.18	-2.57	2.58	1.01	2.61	0.30
PCR	-2.51	1.46	2.73	-3.39	2.62	0.29	2.94	-0.16
RF	-2.13	1.47	3.57	-3.75	2.99	11.59	3.46	-2.71
GBRT	0.79	-1.97	2.74	1.36	1.84	15.31	4.90	4.33
ERT	-1.85	1.81	3.65	-2.95	3.04	16.43	3.20	3.95
NN1	-0.46	1.81	3.65	0.13	3.35	13.30	2.66	2.92
NN2	-1.24	1.79	3.35	-0.97	3.13	5.97	3.22	1.66
NN3	-1.06	2.43	4.84	-2.77	2.63	6.65	4.47	1.80
NN4	-1.03	1.49	2.24	-1.93	2.30	3.35	2.53	1.08
NN5	-0.89	1.67	2.09	-1.56	2.03	4.80	-3.39	1.06

Notes: This table reports monthly R^2_{oos} for the entire panel of investment holding companies, by calendar year. All the numbers are expressed as a percentage.

Table F.3: Monthly predictive R_{oos}^2, by year (Banking)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-20.79	-26.31	-22.88	-8.67	-4.80	-4.94	-1.99	-11.17	-11.61
OLS-2	-0.14	-12.35	-3.89	-0.37	0.64	-1.39	2.81	5.44	1.43
OLS-3	-0.08	-12.24	-3.73	-0.42	0.63	-1.29	2.66	5.48	1.41
LASSO	0.14	-12.33	-3.73	-0.37	0.63	-1.35	2.69	5.40	1.26
ENet	0.57	-12.25	-3.69	-0.33	0.75	-1.22	2.80	5.54	1.71
PCR	-0.05	-12.55	-3.94	-0.40	0.66	-1.25	2.38	5.36	1.44
RF	-0.12	-12.92	-3.83	0.80	0.77	-1.76	3.10	-51.15	1.05
GBRT	0.24	-9.30	-3.63	-0.33	1.27	-0.73	0.89	7.74	-0.12
ERT	0.60	-11.70	-3.12	-1.35	0.70	0.15	2.53	5.98	1.47
NN1	1.34	-9.11	-1.06	-0.44	1.29	-1.96	4.28	7.49	1.81
NN2	0.46	-8.98	-3.97	-0.59	0.80	-2.00	3.82	6.58	1.35
NN3	0.61	-10.11	-4.31	-0.77	0.77	-1.46	3.59	6.92	1.45
NN4	0.82	-12.20	-4.05	-0.74	0.75	-2.32	3.96	6.31	1.66
NN5	0.27	-13.82	-4.09	-0.61	0.75	-2.31	3.66	6.35	1.47

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-5.69	-2.10	-2.41	-9.50	1.12	-9.49	0.15	-9.76
OLS-2	2.33	5.91	2.07	-5.59	3.39	-0.87	6.38	-0.42
OLS-3	2.50	5.73	1.86	-5.74	3.22	-0.96	6.35	-0.43
LASSO	2.21	5.96	1.89	-5.70	3.47	-0.86	6.37	-0.41
ENet	2.62	5.96	2.22	-5.64	3.88	-0.71	6.35	-0.29
PCR	2.07	5.94	1.82	-5.51	3.29	-0.98	6.13	-0.48
RF	2.63	7.28	1.51	-8.82	4.16	0.10	-0.03	-2.34
GBRT	5.41	7.92	-0.70	-3.54	4.87	1.14	6.72	0.12
ERT	3.30	7.02	2.08	-3.22	4.35	8.35	6.83	0.40
NN1	2.07	4.76	1.42	-2.67	4.16	9.52	5.94	1.08
NN2	1.93	4.07	1.38	-1.41	3.69	6.92	5.01	0.25
NN3	2.20	5.92	0.90	-3.63	3.20	9.66	5.68	0.34
NN4	1.03	-1.58	1.06	-3.23	1.61	-0.39	5.14	-0.73
NN5	1.31	-2.25	0.44	-3.15	-1.32	8.22	4.48	-0.34

Notes: This table reports monthly R_{oos}^2 for the entire panel of banks, by calendar year. All the numbers are expressed as a percentage.

Table F.4: Monthly predictive R^2_{oos}, by year (Other Financial Institutions)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-18.74	-21.61	-44.07	-8.04	-1.60	0.45	-3.32	-27.39	-12.13
OLS-2	1.66	-2.90	-1.78	1.09	1.60	-2.15	1.55	4.13	-0.10
OLS-3	1.91	-2.53	-1.70	0.69	1.66	-1.89	1.45	4.16	-0.20
LASSO	1.60	-2.82	-1.78	0.60	1.80	-1.54	1.50	4.36	0.02
ENet	1.63	-2.57	1.50	0.31	2.17	-1.17	1.59	4.63	-0.03
PCR	1.64	-2.63	-1.62	0.77	1.59	-1.79	1.28	4.08	-0.04
RF	1.56	-2.41	-4.73	-1.61	1.41	0.46	-2.38	5.27	-0.23
GBRT	1.44	-1.98	4.34	-2.32	6.19	0.26	2.47	5.80	-1.40
ERT	1.66	-2.00	2.56	-0.41	2.31	-0.63	1.64	4.78	-0.09
NN1	1.75	-2.09	6.96	-0.38	3.51	1.23	1.86	4.41	0.06
NN2	1.41	-1.61	3.40	0.14	2.41	-0.08	1.32	3.79	0.15
NN3	1.53	-2.18	-0.22	0.68	1.59	-0.90	1.65	4.53	0.19
NN4	1.29	-2.38	-1.56	1.24	1.57	-2.23	1.64	4.04	0.11
NN5	1.19	-2.43	-1.26	1.32	1.72	-2.29	1.58	4.10	0.16

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-1.56	-5.55	-1.97	-3.67	-4.01	-13.69	1.53	-12.51
OLS-2	-0.36	1.19	0.66	-3.23	1.60	0.62	1.55	0.21
OLS-3	-0.31	1.12	0.76	-3.24	1.60	0.56	1.56	0.19
LASSO	-0.29	1.26	0.75	-2.78	1.78	2.31	1.50	0.35
ENet	-0.18	1.45	0.79	-2.28	1.85	4.93	1.48	1.10
PCR	-0.28	1.11	0.67	-3.09	1.61	0.61	1.56	0.20
RF	-0.32	-0.18	-1.36	-2853.44	2.06	4.06	-5.43	-88.06
GBRT	0.54	2.57	0.79	1.37	2.19	5.08	1.89	1.67
ERT	-0.29	1.39	0.79	-15.88	2.00	8.68	1.68	1.10
NN1	-0.09	1.50	0.79	-0.04	2.06	8.82	1.34	2.38
NN2	-0.07	0.71	0.45	0.23	0.97	6.68	1.31	1.53
NN3	0.01	0.58	-0.10	0.13	1.09	2.18	1.09	0.62
NN4	-0.17	0.43	0.17	0.78	0.82	3.21	0.87	0.51
NN5	-0.15	0.08	0.37	1.51	-0.05	7.01	-0.52	0.81

Notes: This table reports monthly R^2_{oos} for the entire panel of financial institutions other than banks, by calendar year. All the numbers are expressed as a percentage.

Table F.5: Monthly predictive R^2_{oos}, by year (Agriculture)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-2476.29	-2445.86	-2865.06	-455.63	-328.18	-265.52	-425.48	-686.18	-459.60
OLS-2	-0.45	-1.60	-2.62	0.90	0.25	-3.37	0.07	2.05	-2.80
OLS-3	-0.53	-1.58	-2.67	0.87	0.27	-3.44	0.06	2.05	-2.82
LASSO	-0.65	-0.02	-2.95	1.05	0.64	-2.52	0.58	1.68	-1.14
ENet	-1.21	0.46	1.34	0.59	1.58	-1.63	1.36	1.92	-1.85
PCR	-1.55	-1.95	-2.46	1.04	0.46	-3.39	-0.01	2.75	-4.13
RF	-28.91	-11.89	-10.49	-0.92	-2.18	-2.39	0.63	0.50	-0.92
GBRT	-19.68	-12.06	-1.15	-1.69	-0.28	-1.37	1.34	1.29	-4.89
ERT	-2.71	-2.32	-4.46	-1.91	0.15	-2.87	0.59	1.74	-2.04
NN1	-0.12	-1.60	-2.44	2.15	1.03	-1.59	1.18	1.54	-5.29
NN2	0.19	0.27	-1.16	0.91	0.47	-1.24	0.86	1.08	-2.86
NN3	0.25	0.35	0.84	0.69	0.35	-1.03	0.10	0.89	-0.69
NN4	-1.45	-1.55	2.02	-0.58	-0.76	1.33	0.31	-1.14	0.12
NN5	-3.65	-3.45	2.39	-1.06	-0.66	2.13	-0.94	-1.32	0.16

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-301.56	-473.91	-339.92	-512.46	-38.58	-1680.19	-399.71	-783.31
OLS-2	-1.34	0.67	-0.66	-3.63	0.62	0.34	1.31	-0.40
OLS-3	-1.34	0.67	-0.72	-3.59	0.59	0.34	1.30	-0.42
LASSO	-0.85	1.02	-0.08	-2.82	0.48	1.82	0.95	-0.10
ENet	-0.52	2.32	-0.18	-1.93	1.06	9.71	1.44	0.75
PCR	-0.85	0.63	-1.16	-3.95	0.64	0.01	0.94	-0.42
RF	-0.52	0.39	-0.33	-1.40	0.33	1.87	0.68	-2.87
GBRT	-2.11	0.10	0.41	-0.03	1.35	2.61	0.62	-1.60
ERT	-0.96	1.21	0.21	-2.11	0.83	10.84	1.56	-0.77
NN1	-2.95	2.40	0.01	-2.26	1.03	3.26	1.91	0.16
NN2	-1.85	1.56	0.09	-2.87	0.75	1.86	1.14	0.08
NN3	-0.79	0.87	0.12	-2.71	0.39	0.53	0.90	0.23
NN4	0.05	-0.37	-0.04	1.42	0.11	-0.10	-0.61	-0.01
NN5	0.36	-0.82	-0.83	0.07	-0.44	-0.06	-1.08	-0.37

Notes: This table reports monthly R^2_{oos} for the entire panel of financial institutions other than banks, by calendar year. All the numbers are expressed as a percentage.

Table F.6: Monthly predictive R^2_{oos}, by year (Mining)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-17.65	-15.34	-7.04	-11.84	1.99	-3.52	-7.51	-6.19	-13.96
OLS-2	0.70	-0.34	-1.87	1.35	2.35	-1.39	-1.31	-1.16	-2.17
OLS-3	0.66	-0.09	-1.93	1.12	2.43	-1.40	-1.28	-0.88	-2.10
LASSO	0.62	0.49	-1.85	-0.59	3.11	-0.46	-1.14	-0.81	-2.21
ENet	0.64	0.59	2.72	-0.65	3.36	-0.41	-1.20	-0.85	-2.27
PCR	0.47	-0.04	-1.97	1.03	2.23	-1.18	-1.08	-0.64	-2.07
RF	0.40	-0.21	-0.06	-2.20	2.90	-0.76	-6.17	-0.99	-2.59
GBRT	0.25	-1.79	12.94	-4.24	6.24	0.25	-9.38	0.78	-3.40
ERT	0.75	0.37	3.15	-1.42	2.62	-0.19	-1.20	-0.87	-2.37
NN1	0.03	0.09	3.73	-0.02	3.92	-0.22	-1.12	-0.92	-1.92
NN2	0.06	0.11	1.08	-0.31	1.18	-0.08	-0.04	0.09	-2.17
NN3	0.36	0.17	0.21	-0.03	1.56	-0.39	-0.64	-0.33	-2.16
NN4	0.27	0.39	-1.31	1.07	1.85	-1.23	-1.07	-0.09	-1.29
NN5	0.44	0.36	-1.37	1.95	1.95	-1.21	-0.87	-0.23	-1.18

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-1.08	-8.87	-10.13	0.83	-4.28	-12.03	-2.69	-7.58
OLS-2	-1.22	0.69	-0.72	-1.05	-0.13	0.23	0.74	-0.24
OLS-3	-1.10	0.67	-0.62	-1.02	0.08	0.24	0.68	-0.22
LASSO	-1.08	0.87	-0.75	-0.35	0.07	2.81	0.61	0.11
ENet	-1.10	0.92	-0.81	-0.28	0.08	3.17	0.64	0.63
PCR	-1.19	0.76	-0.54	-0.66	0.06	0.15	0.19	-0.24
RF	-0.39	-0.48	0.20	-0.10	0.20	3.57	0.99	-0.05
GBRT	1.63	-2.74	0.26	2.37	-0.45	2.04	-1.27	0.82
ERT	-1.29	0.78	-0.71	-1.09	-0.02	6.14	0.74	0.90
NN1	-3.06	1.36	-2.30	-0.71	-0.08	5.68	-1.97	0.80
NN2	-3.64	1.89	-1.37	-1.14	-0.57	2.09	-1.03	-0.01
NN3	-2.52	1.77	-2.25	-1.59	-0.18	0.69	-4.10	-0.40
NN4	-4.01	1.24	-0.23	-1.31	0.28	0.73	-0.18	-0.28
NN5	-0.04	1.25	-0.20	-0.86	-0.14	0.26	-3.35	-0.06

Notes: This table reports monthly R^2_{oos} for the entire panel of mining firms, by calendar year. All the numbers are expressed as a percentage.

Table F.7: Monthly predictive R^2_{oos}, by year (Construction)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-179.17	-81.44	-321.64	-12.56	-83.28	-74.92	-23.51	-133.44	-69.04
OLS-2	0.09	-2.66	-0.98	0.12	0.29	-2.47	1.16	1.31	-1.81
OLS-3	0.63	-0.64	-1.24	-0.01	0.19	-2.28	1.36	0.84	-2.20
LASSO	-0.26	-1.41	2.02	-0.56	0.77	-0.77	1.33	1.36	-1.41
ENet	0.62	-0.29	9.19	-0.76	1.03	-0.60	1.52	1.62	-1.81
PCR	0.44	-0.25	-0.01	-0.32	-0.43	-1.15	-0.01	1.05	-1.39
RF	-2.08	-16.56	-123.19	-0.58	1.12	-0.13	0.92	1.06	-3.03
GBRT	-5.79	-7.69	1.59	-0.79	-0.16	-1.63	0.56	0.95	-6.83
ERT	-0.29	-1.52	1.08	-0.91	0.77	-0.49	0.96	1.08	-1.34
NN1	0.81	-1.45	6.76	0.53	2.18	-1.93	1.74	2.78	1.00
NN2	0.02	-0.34	1.21	0.05	0.52	-0.63	0.71	0.87	-0.09
NN3	-0.07	0.69	1.55	-0.06	-0.26	0.93	0.32	0.64	-0.12
NN4	-0.13	0.80	0.65	0.03	-0.34	0.03	0.47	-1.04	-0.19
NN5	-0.34	0.93	0.61	-0.10	-0.11	0.54	-0.66	-0.14	-0.23

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-79.52	-66.86	-58.51	-26.98	-11.59	-144.31	-31.74	-51.99
OLS-2	-1.27	1.18	0.55	-2.85	0.71	1.31	1.39	0.02
OLS-3	-0.77	0.76	0.01	-2.92	0.60	1.44	1.23	-0.03
LASSO	-0.54	1.20	1.55	-2.22	0.70	3.88	1.47	0.05
ENet	-0.60	1.10	1.68	-1.91	0.85	6.40	1.29	0.52
PCR	0.38	0.32	1.19	-2.78	0.34	0.41	1.15	-0.16
RF	-0.70	0.99	1.18	-1.40	0.82	2.14	-47.79	-9.04
GBRT	-1.45	0.68	-0.06	1.50	1.00	5.95	2.20	-0.41
ERT	-0.58	1.00	1.32	-2.03	0.71	9.26	1.39	0.03
NN1	-0.74	-0.57	1.11	-2.62	1.08	4.52	1.04	1.04
NN2	-0.33	0.49	0.65	-1.85	0.56	1.20	0.53	0.21
NN3	-0.26	0.22	0.93	-0.33	0.31	0.75	0.87	0.19
NN4	-0.32	0.17	0.82	-1.20	0.35	0.49	0.20	0.09
NN5	-0.12	0.68	0.74	-1.46	0.41	0.73	0.59	0.05

Notes: This table reports monthly R^2_{oos} for the entire panel of construction firms, by calendar year. All the numbers are expressed as a percentage.

Table F.8: Monthly predictive R^2_{oos}, by year (Other Manufacturing)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-4.63	-2.87	-6.64	-3.73	2.36	2.27	0.25	-11.31	-6.33
OLS-2	0.78	-1.08	-2.44	1.29	1.52	-1.57	1.05	2.79	-0.52
OLS-3	0.78	-0.89	-2.38	1.01	1.53	-1.47	1.05	2.81	-0.52
LASSO	0.79	-0.68	-2.31	0.61	1.84	-0.49	1.17	3.09	-0.36
ENet	0.88	-0.37	3.59	0.35	2.16	-0.03	1.26	3.32	-0.44
PCR	0.67	-1.36	-2.67	1.18	1.54	-1.45	0.96	2.87	-0.43
RF	0.87	2.17	-4.33	-3.71	4.11	1.89	1.38	3.84	-0.65
GBRT	0.77	2.07	4.11	-2.65	5.47	1.47	1.43	2.31	-1.58
ERT	0.87	0.27	4.73	-0.72	2.55	0.89	1.24	3.37	-0.45
NN1	0.65	-1.15	6.93	0.63	3.14	0.85	1.51	3.90	-0.41
NN2	-0.47	0.10	10.31	-0.37	3.89	2.31	1.45	3.43	-0.37
NN3	-0.36	-0.36	11.11	1.76	2.73	1.58	1.49	3.82	-0.6
NN4	-1.02	-0.87	2.47	0.04	3.95	0.36	1.33	1.34	-1.47
NN5	-2.56	-3.18	8.28	-1.34	-0.74	0.14	1.27	4.43	-0.04

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-1.95	-3.90	-1.19	1.54	-0.07	-10.03	-0.14	-3.66
OLS-2	-1.29	0.98	0.74	-1.21	0.61	0.69	1.02	0.23
OLS-3	-1.16	0.90	0.76	-1.13	0.59	0.66	0.92	0.20
LASSO	-1.01	1.09	0.88	-0.70	0.90	2.32	0.98	0.54
ENet	-0.92	1.19	0.93	-0.38	0.95	3.18	0.97	1.33
PCR	-1.46	0.99	0.66	-1.14	0.65	0.67	0.96	0.17
RF	-0.05	2.25	0.99	2.54	1.73	0.85	-2.49	-0.01
GBRT	-1.09	1.73	1.07	2.13	1.26	2.81	0.98	1.34
ERT	-1.11	1.26	0.92	-0.12	1.02	3.53	1.08	1.44
NN1	-1.00	1.14	-0.55	1.44	1.26	3.26	0.72	1.83
NN2	-0.13	1.65	0.52	1.36	0.68	3.82	-0.87	2.19
NN3	-0.11	1.82	-0.64	1.80	1.06	3.65	-0.36	2.49
NN4	0.85	-1.69	-0.87	1.66	-4.13	0.67	-4.47	0.14
NN5	-0.01	-10.51	-6.44	-6.09	1.16	-2.24	0.44	-0.84

Notes: This table reports monthly R^2_{oos} for the entire panel of other manufacturing firms, by calendar year. All the numbers are expressed as a percentage.

Table F.9: Monthly predictive R^2_{oos}, by year (Chemicals)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-5.52	-8.57	-43.54	-2.74	-0.94	0.41	0.15	-10.98	-2.88
OLS-2	0.61	-1.28	-1.90	0.94	0.95	-0.94	0.75	1.94	0.03
OLS-3	0.56	-1.26	-1.87	0.86	0.95	-0.93	0.74	1.96	0.04
LASSO	0.45	-0.89	-1.72	0.37	1.19	-0.46	0.80	1.94	0.24
ENet	0.55	-0.96	0.97	0.09	1.46	-0.32	0.97	2.17	0.17
PCR	0.50	-1.38	-2.05	0.47	1.04	-0.76	0.82	2.02	0.01
RF	0.96	-0.16	-8.65	-2.68	1.69	0.69	1.05	2.17	0.05
GBRT	1.04	-0.78	1.40	-1.46	2.33	1.02	1.50	2.92	-0.73
ERT	0.48	-0.63	1.81	-0.74	1.31	-0.20	0.86	2.06	0.22
NN1	-2.00	-4.89	3.89	-0.34	1.76	-0.98	0.97	2.34	-0.19
NN2	-0.06	-1.45	1.99	0.47	1.60	-0.15	0.89	3.10	0.13
NN3	0.39	-1.66	0.16	0.74	1.28	-0.16	0.83	3.22	0.13
NN4	0.44	-1.48	-1.25	1.51	-0.09	-1.03	0.73	0.06	0.08
NN5	0.44	-1.10	-1.44	-0.11	0.44	-0.61	0.74	-0.44	0.13

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-0.36	-2.35	-0.45	-1.88	-0.02	-6.52	-2.39	-5.59
OLS-2	0.00	-0.53	0.53	-1.27	0.43	0.75	-0.44	0.08
OLS-3	-0.01	-0.54	0.53	-1.25	0.42	0.74	-0.44	0.07
LASSO	0.02	-0.47	0.65	-0.99	0.48	1.06	-0.37	0.15
ENet	0.13	-0.48	0.64	-0.85	0.52	1.65	-0.45	0.42
PCR	-0.05	-0.40	0.49	-1.23	0.40	0.66	-0.37	0.01
RF	0.00	-15.42	0.69	-0.33	-4.46	0.34	-0.46	-2.10
GBRT	0.71	-1.51	0.60	1.27	-9.34	1.30	-1.70	-0.47
ERT	0.07	-0.43	0.67	-0.27	0.51	1.49	-0.40	0.39
NN1	0.49	0.07	0.27	0.58	0.75	2.61	0.41	0.63
NN2	0.39	-0.13	0.14	-0.33	0.57	1.43	0.33	0.59
NN3	0.37	-0.99	0.19	-0.27	0.48	1.08	0.16	0.39
NN4	0.13	-0.13	0.26	-1.35	0.36	0.30	0.01	0.06
NN5	0.24	-0.28	0.76	-0.47	0.48	0.23	-0.47	-0.06

Notes: This table reports monthly R^2_{oos} for the entire panel of chemical firms, by calendar year. All the numbers are expressed as a percentage.

Table F.10: Monthly predictive R^2_{oos}, by year (IT)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-10.09	-563.63	-48.12	-4.34	0.57	2.29	-0.74	-14.32	-9.56
OLS-2	0.12	-1.25	-4.10	1.66	1.50	-3.05	0.13	2.90	-0.14
OLS-3	0.07	-1.20	-4.08	1.58	1.50	-3.02	0.14	2.92	-0.14
LASSO	0.02	-0.92	-3.19	0.81	2.05	-1.98	0.15	2.92	-0.01
ENet	0.08	-0.86	3.06	0.02	2.73	-1.54	0.30	3.24	-0.07
PCR	0.05	-1.41	-4.39	1.53	1.61	-2.79	0.21	2.68	-0.08
RF	-1.00	-1.03	-41.04	-4.83	5.02	0.05	1.06	4.21	-0.57
GBRT	0.53	-1.66	-2.25	-4.21	6.78	1.98	0.69	2.69	-0.90
ERT	-0.14	-0.22	3.64	-1.53	2.80	-0.90	0.23	3.18	-0.07
NN1	0.16	-0.71	11.27	1.86	2.68	-0.24	0.55	4.14	-0.04
NN2	-0.35	-0.43	4.95	1.18	1.83	-1.55	0.26	3.09	-1.72
NN3	-0.26	-0.91	-0.76	1.95	1.86	0.47	0.40	4.95	0.09
NN4	-0.55	-0.69	0.40	-0.67	2.11	0.47	0.36	3.15	-0.52
NN5	0.21	-0.72	0.28	1.40	-7.31	2.84	0.36	1.85	0.04

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-5.95	-4.03	-0.83	-7.00	-0.27	-16.15	-1.34	-37.34
OLS-2	-0.95	0.55	1.23	-3.11	0.93	1.46	0.40	0.08
OLS-3	-0.94	0.53	1.22	-3.06	0.90	1.45	0.40	0.07
LASSO	-0.53	0.76	1.22	-2.02	1.33	2.63	0.34	0.30
ENet	-0.47	0.92	1.29	-1.53	1.32	2.88	0.33	0.87
PCR	-0.90	0.66	1.04	-2.86	1.20	1.27	0.39	0.04
RF	-0.57	0.61	1.42	1.45	1.89	0.09	-0.19	-3.87
GBRT	-1.23	0.81	1.14	-3.26	1.30	2.83	0.39	-0.02
ERT	-0.60	0.90	1.25	-1.09	1.48	2.71	0.36	0.74
NN1	-1.48	0.70	0.34	2.19	0.89	2.21	0.31	1.92
NN2	-0.31	0.67	0.25	-0.62	1.07	1.33	0.06	0.88
NN3	-0.59	0.97	1.03	0.73	0.67	-0.03	0.00	0.61
NN4	-0.61	-0.71	0.81	2.05	-2.56	-1.63	-0.67	-0.22
NN5	-0.03	-1.48	1.00	2.03	-8.15	-1.41	-0.64	-0.49

Notes: This table reports monthly R^2_{oos} for the entire panel of IT firms, by calendar year. All the numbers are expressed as a percentage.

Table F.11: Monthly predictive R^2_{oos}, by year (Transportation)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-28.92	-36.46	-89.65	-15.84	-5.24	1.03	-3.19	-30.66	-17.52
OLS-2	0.71	-0.30	-2.02	0.71	0.58	-1.36	0.28	1.66	-0.93
OLS-3	0.87	0.31	-2.09	0.41	0.88	-1.00	0.27	1.49	-0.42
LASSO	0.78	-0.27	-1.96	0.07	1.50	-0.28	0.46	2.06	-0.94
ENet	0.83	0.35	5.45	-0.03	1.88	-0.26	0.53	2.25	-0.95
PCR	0.11	-1.10	-1.49	0.36	0.91	-1.10	0.25	1.52	0.19
RF	0.84	1.40	-25.66	-1.31	2.13	-0.15	1.00	2.78	-25.28
GBRT	0.26	2.00	-3.19	0.04	4.55	-0.69	-2.00	1.11	-6.60
ERT	0.78	0.83	3.15	-0.46	1.57	-0.16	0.37	1.99	-1.01
NN1	0.67	-0.07	5.88	1.22	2.03	-2.25	0.11	2.18	1.37
NN2	0.42	-0.08	0.56	0.98	1.05	-0.61	0.26	1.98	-0.16
NN3	0.49	0.00	-0.84	1.05	0.78	-0.32	0.35	2.01	-0.22
NN4	0.57	-0.04	-1.11	1.09	0.90	-1.17	0.36	1.84	-0.11
NN5	0.30	-0.06	-0.74	1.05	0.94	-1.54	0.34	1.88	-0.11

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-14.85	-5.93	-8.22	-4.87	-15.75	-34.96	-6.06	-21.57
OLS-2	-2.06	0.20	0.06	-1.41	0.82	-0.10	1.15	-0.13
OLS-3	-1.33	0.15	0.34	-1.52	0.97	-0.24	1.05	-0.07
LASSO	-1.75	0.41	0.04	-0.59	1.05	5.05	1.33	0.52
ENet	-1.68	0.45	0.13	-0.18	1.12	6.13	1.29	1.38
PCR	-1.29	0.06	0.11	-1.34	0.45	0.07	1.13	-0.09
RF	-2.35	0.44	-122.72	-19.05	1.03	4.11	1.84	-10.76
GBRT	-1.22	-0.87	0.38	4.65	-8.15	5.42	0.28	0.00
ERT	-2.13	0.41	-0.31	0.07	1.10	5.67	1.54	1.03
NN1	-0.33	0.20	0.30	1.74	2.23	4.81	-3.61	1.32
NN2	-2.96	0.20	0.17	0.93	0.82	4.97	-0.49	0.72
NN3	-1.49	0.23	0.07	-0.24	0.54	1.47	0.43	0.34
NN4	-1.92	0.18	0.09	-0.19	0.55	0.77	0.49	0.18
NN5	-1.42	0.05	0.09	-0.96	0.80	0.43	0.57	0.13

Notes: This table reports monthly R^2_{oos} for the entire panel of transportation firms, by calendar year. All the numbers are expressed as a percentage.

Table F.12: Monthly predictive R^2_{oos}, by year (Utilities)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-17.34	-20.53	-52.7	-8.15	-3.89	-15.17	-6.76	-18.22	-22.26
OLS-2	1.64	-0.50	-2.31	1.27	1.38	-0.67	0.61	3.07	-0.09
OLS-3	1.62	-0.25	-2.08	0.76	1.53	-0.56	0.66	3.36	-0.19
LASSO	1.61	-0.19	-2.12	0.62	1.59	0.30	0.75	2.82	0.29
ENet	1.75	0.15	2.43	0.26	1.95	0.88	0.75	3.04	0.30
PCR	1.59	-0.44	-2.37	1.37	1.78	-0.65	0.43	3.40	-0.64
RF	1.36	0.52	1.47	-6.10	4.79	1.93	0.76	4.06	-0.24
GBRT	-1.75	3.12	3.60	-4.29	7.41	1.44	-0.36	5.98	-2.52
ERT	1.77	1.13	4.98	-0.96	2.85	1.54	0.81	3.36	0.12
NN1	0.82	0.33	11.63	-0.04	2.81	-0.15	0.95	3.06	0.34
NN2	1.91	0.21	2.68	0.07	1.23	-0.20	0.79	2.51	0.07
NN3	1.88	-0.17	-0.47	0.86	1.40	-0.21	0.74	3.46	0.25
NN4	1.49	-0.14	-1.34	0.92	1.23	-0.21	0.78	2.27	0.14
NN5	1.59	-0.03	-1.12	1.01	1.16	-0.22	0.76	2.75	0.17

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-8.76	-5.35	-4.74	-2.66	-0.37	-6.98	-15.83	-14.11
OLS-2	-1.30	1.16	0.82	-1.42	0.54	0.22	1.57	0.26
OLS-3	-1.00	0.97	0.72	-1.50	0.64	0.16	1.40	0.23
LASSO	-1.17	1.33	0.83	-0.90	0.83	0.50	1.39	0.36
ENet	-1.03	1.52	0.87	-0.36	0.89	1.04	1.40	1.04
PCR	-1.47	1.31	0.60	-1.6	0.37	0.21	1.51	0.26
RF	-0.57	1.64	0.90	3.32	1.91	0.01	-105.37	-3.42
GBRT	-2.04	1.70	0.46	2.19	1.65	1.55	1.88	0.97
ERT	-1.13	1.63	0.88	0.22	1.10	1.85	1.48	1.50
NN1	-0.66	1.84	0.57	-0.11	0.81	1.79	1.81	2.20
NN2	-0.23	0.21	0.38	0.05	0.76	0.96	-0.14	0.85
NN3	-0.26	0.41	0.40	0.13	0.47	0.45	0.41	0.55
NN4	-0.18	0.25	-0.01	0.13	0.53	1.07	0.13	0.51
NN5	-0.64	-0.09	0.46	0.50	0.61	0.13	-0.93	0.31

Notes: This table reports monthly R^2_{oos} for the entire panel of utility firms, by calendar year. All the numbers are expressed as a percentage.

Table F.13: Monthly predictive R_{oos}^2, by year (Wholesale)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-24.97	-44.46	-77.32	-9.81	-2.26	-5.96	-6.95	-27.71	-32.47
OLS-2	0.43	-1.29	-2.58	1.06	1.47	-2.21	0.80	2.11	-0.97
OLS-3	0.45	-1.20	-2.68	0.90	1.40	-2.12	0.83	2.15	-1.03
LASSO	0.61	-0.61	-0.19	0.41	2.05	-0.58	0.84	2.41	-0.58
ENet	0.72	-0.08	7.02	0.42	2.78	-0.59	0.99	2.70	-0.85
PCR	0.41	-1.46	-2.86	0.95	1.70	-1.97	0.74	2.20	-1.02
RF	-0.20	-15.91	3.69	-3.22	3.33	-0.37	1.10	3.11	-1.08
GBRT	-1.65	-12.41	4.38	-3.00	5.97	-0.61	1.68	3.32	-1.32
ERT	0.42	-2.07	2.29	-0.79	2.40	-0.42	0.96	2.75	-0.79
NN1	0.46	-1.13	6.51	2.10	3.50	-2.30	0.81	2.77	-0.51
NN2	0.15	-0.10	0.79	0.99	1.74	-0.38	0.71	1.29	-0.18
NN3	0.21	-0.04	0.66	0.91	1.39	-0.40	0.50	1.69	-0.21
NN4	0.34	-0.39	-1.08	0.93	1.54	-1.31	0.84	2.08	-0.12
NN5	0.33	-0.45	-1.42	0.91	1.68	-1.27	0.84	2.27	-0.13

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-14.55	-7.48	-5.54	2.11	-5.40	-31.4	-5.09	-18.7
OLS-2	-1.46	1.49	0.17	-0.45	0.55	0.66	1.39	0.22
OLS-3	-1.28	1.42	0.30	-0.36	0.46	0.63	1.49	0.20
LASSO	-1.07	1.42	0.23	-0.20	0.74	4.02	1.21	0.76
ENet	-0.94	1.60	0.21	0.15	0.78	5.92	1.23	1.64
PCR	-1.51	1.38	0.08	-0.28	0.53	0.69	1.40	0.19
RF	-1.35	1.65	0.22	-0.11	-2.72	1.02	-2.31	-0.89
GBRT	-1.28	1.82	0.20	1.80	1.39	4.37	1.29	0.39
ERT	-1.28	1.48	0.29	0.00	0.86	5.55	1.37	0.86
NN1	-1.21	2.29	0.21	0.33	0.80	7.31	0.77	1.96
NN2	-1.06	1.03	0.37	-0.12	0.68	3.00	0.49	0.75
NN3	-0.54	0.96	0.28	-0.27	0.68	0.77	1.10	0.58
NN4	-1.31	0.97	0.17	-0.31	0.69	0.55	0.79	0.36
NN5	-1.88	0.89	0.26	-0.46	0.64	0.18	0.70	0.26

Notes: This table reports monthly R_{oos}^2 for the entire panel of wholesale trade firms, by calendar year. All the numbers are expressed as a percentage.

Table F.14: Monthly predictive R^2_{oos}, by year (Retail)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-16.64	-24.09	-31.08	-2.22	-1.96	-0.35	-1.25	-27.47	-9.09
OLS-2	1.05	-1.85	-1.09	0.49	1.18	-0.15	0.80	2.00	-0.13
OLS-3	0.80	-1.39	-1.07	-0.08	1.09	-0.09	0.63	2.12	-0.02
LASSO	0.81	-1.72	-1.04	-0.64	1.59	0.39	0.87	2.19	-0.11
ENet	0.98	-1.16	5.46	-0.95	2.23	0.93	1.08	2.52	-0.11
PCR	0.83	-1.75	-0.77	0.10	1.07	-0.32	0.87	1.86	0.22
RF	0.80	-0.16	3.52	-9.83	3.73	0.04	0.79	1.67	-9.30
GBRT	-1.51	-5.36	5.24	-6.18	6.86	-1.03	1.56	4.06	-73.94
ERT	1.05	-0.99	5.67	-2.64	1.69	0.37	0.92	2.25	-0.11
NN1	0.15	0.01	8.05	-2.16	3.03	3.23	0.89	1.78	0.21
NN2	-0.08	-0.66	-0.07	-1.17	2.17	2.25	0.85	1.17	-0.04
NN3	0.06	-0.64	-0.44	-0.48	1.73	0.94	0.54	2.79	0.12
NN4	0.02	-0.72	-0.62	0.69	1.13	-0.05	1.15	1.32	-0.08
NN5	-0.17	-1.11	-0.46	0.68	1.06	-0.05	0.67	1.61	0.18

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-9.83	-4.18	-5.97	-0.38	-3.45	-16.3	1.66	-7.87
OLS-2	-1.28	0.39	-0.12	-0.79	0.39	0.57	0.34	0.17
OLS-3	-1.23	0.21	-0.01	-0.72	0.27	0.48	0.39	0.09
LASSO	-1.10	0.40	-0.16	-0.36	0.43	2.16	0.30	0.24
ENet	-1.10	0.49	-0.24	0.24	0.45	3.59	0.29	1.02
PCR	-1.23	0.30	-0.20	-0.90	0.28	0.18	0.34	0.08
RF	-1.38	0.56	-0.16	0.91	0.31	5.68	0.48	-0.49
GBRT	-0.38	0.96	-0.59	3.74	-0.72	5.88	0.33	-2.24
ERT	-1.29	0.36	-0.18	-0.43	0.46	6.42	0.30	1.03
NN1	-1.35	0.05	-0.56	-0.40	0.87	5.22	0.37	1.34
NN2	-3.46	0.50	-0.10	-0.36	0.69	3.41	0.23	0.39
NN3	-1.54	0.19	-0.28	-0.31	0.25	4.33	-0.01	0.51
NN4	-0.58	-0.67	-0.25	-0.50	-1.28	1.29	0.17	0.19
NN5	0.05	-0.54	-0.03	-0.43	-1.37	-0.47	0.24	0.03

Notes: This table reports monthly R^2_{oos} for the entire panel of retail trade firms, by calendar year. All the numbers are expressed as a percentage.

Table F.15: Monthly predictive R_{oos}^2, by year (Services)

	2006	2007	2008	2009	2010	2011	2012	2013	2014
OLS	-11.6	-11.27	-34.6	-3.13	1.62	1.64	0.17	-14.72	-8.80
OLS-2	0.43	-1.44	-2.94	1.67	1.24	-1.52	0.65	2.72	-0.54
OLS-3	0.33	-1.22	-2.87	1.28	1.26	-1.41	0.66	2.75	-0.68
LASSO	0.58	-0.60	0.41	0.57	1.59	-0.32	0.89	3.09	-0.41
ENet	0.76	-0.68	4.18	0.64	2.06	-0.14	0.95	3.32	-0.47
PCR	0.42	-1.51	-3.19	1.49	1.25	-1.47	0.67	2.73	-0.51
RF	0.32	1.02	4.98	-4.77	5.03	1.08	0.68	4.71	-1.34
GBRT	-1.64	1.57	3.75	-3.63	6.35	-0.48	-0.17	6.25	-3.54
ERT	0.54	0.41	4.74	-1.41	2.79	0.94	0.93	3.58	-0.6
NN1	-2.74	-4.09	7.22	0.45	2.28	-0.3	-0.52	1.95	-1.62
NN2	-0.08	-13.11	10.9	-0.28	2.08	0.00	-0.03	2.05	-3.82
NN3	-0.67	-18.94	9.90	2.14	1.29	-1.21	-0.62	3.38	-0.75
NN4	-3.68	-23.36	8.67	-1.87	1.99	-1.55	-3.75	4.43	-1.12
NN5	-4.18	-13.10	5.27	0.57	2.10	-7.83	-2.07	0.62	-0.61

	2015	2016	2017	2018	2019	2020	2021	All Years
OLS	-0.82	-1.68	-2.33	0.47	-1.61	-11.60	0.53	-6.90
OLS-2	-0.99	0.46	1.05	-1.71	0.58	1.03	0.52	0.16
OLS-3	-0.86	0.34	1.01	-1.60	0.63	0.96	0.55	0.13
LASSO	-0.35	0.51	1.30	-0.51	1.07	3.15	0.37	0.82
ENet	-0.30	0.64	1.36	-0.32	1.11	4.02	0.37	1.34
PCR	-0.97	0.49	0.95	-1.51	0.56	0.86	0.57	0.11
RF	-29.41	1.86	1.60	3.99	1.77	1.26	-0.31	-0.29
GBRT	-0.28	2.36	1.45	4.13	1.19	3.21	0.54	1.23
ERT	-0.11	0.68	1.44	0.78	1.26	3.50	0.44	1.32
NN1	-2.99	0.63	1.36	1.13	0.69	4.83	0.42	1.12
NN2	-5.90	0.46	1.29	0.59	0.71	4.55	-0.24	0.76
NN3	-4.52	0.86	0.90	2.10	0.90	2.57	0.43	0.71
NN4	-2.57	-2.42	-0.16	0.79	0.28	3.79	-0.52	-0.52
NN5	-0.11	0.39	-0.64	-3.85	-3.35	5.57	-1.63	-0.46

Notes: This table reports monthly R_{oos}^2 for the entire panel of services firms, by calendar year. All the numbers are expressed as a percentage.

G Relative variable importance for macroeconomic predictors for the U.S. stock market

Table G.16: Relative variable importance for macroeconomic predictors for general stocks

	PLS	PCR	ENet+H	GLM+H	RF	GBRT+H	NN1	NN2	NN3	NN4	NN5
dp	12.52	14.12	2.49	4.54	5.80	6.05	15.57	17.58	14.84	13.95	13.15
ep	12.25	13.52	3.27	7.37	6.27	2.85	8.86	8.09	7.34	6.54	6.47
bm	14.21	14.83	33.95	43.46	10.94	12.49	28.57	27.18	27.92	26.95	27.90
ntis	11.25	9.10	1.30	4.89	13.02	13.79	18.37	19.26	20.15	19.59	18.68
tbl	14.02	15.29	13.29	7.90	11.98	19.49	17.18	16.40	17.76	20.99	21.06
tms	11.35	10.66	0.31	5.87	16.81	15.27	10.79	10.59	10.91	10.38	10.33
dfy	17.17	15.68	42.13	24.10	24.37	22.93	0.09	0.06	0.06	0.04	0.12
svar	7.22	6.8	3.26	1.87	10.82	7.13	0.57	0.85	1.02	1.57	2.29

Notes: This table reports the variable importance for eight macroeconomic variables for the U.S. stock market, obtained from Gu et al. (2020) to be used as a basis of comparison with Table 8.

H Yearly relative variable importance for REIT-level predictors

Table H.1: Top 10 relative variable importance - OLS

	2006	2007	2008	2009	2010	2011	2012	2013	2014
rd_sale	50%	50%	49%	48%	49%	49%	49%	49%	49%
rd_mve	49%	49%	49%	48%	48%	48%	48%	48%	49%
beta	0%	0%	1%	1%	1%	1%	1%	1%	1%
betasq	0%	0%	1%	1%	1%	1%	1%	1%	1%
securedind	0%	0%	0%	0%	0%	0%	0%	0%	0%
mvel1	0%	0%	0%	0%	0%	0%	0%	0%	0%
dolvol	0%	0%	0%	0%	0%	0%	0%	0%	0%
currat	0%	0%	0%	0%	0%	0%	0%	0%	0%
quick	0%	0%	0%	0%	0%	0%	0%	0%	0%
ill	0%	0%	0%	0%	0%	0%	0%	0%	0%

	2015	2016	2017	2018	2019	2020	2021	All Years
rd_sale	49%	49%	48%	48%	48%	47%	46%	48%
rd_mve	49%	49%	48%	48%	48%	47%	46%	48%
beta	1%	1%	1%	2%	2%	3%	3%	1%
betasq	1%	1%	1%	2%	2%	3%	3%	1%
securedind	0%	0%	0%	0%	0%	0%	0%	0%
mvel1	0%	0%	0%	0%	0%	0%	0%	0%
dolvol	0%	0%	0%	0%	0%	0%	0%	0%
currat	0%	0%	0%	0%	0%	0%	0%	0%
quick	0%	0%	0%	0%	0%	0%	0%	0%
ill	0%	0%	0%	0%	0%	0%	0%	0%

Notes: This table reports the top 10 most influential variables for the OLS model, by year.

Table H.2: Top 10 relative variable importance - ENet

	2006	2007	2008	2009	2010	2011	2012	2013	2014
securedind	92%	100%	100%	96%	99%	99%	99%	100%	100%
mom1m	8%	0%	0%	0%	1%	1%	1%	0%	0%
convind	0%	0%	0%	0%	0%	0%	0%	0%	0%
mom12m	0%	0%	0%	3%	0%	0%	0%	0%	0%
lev	0%	0%	0%	0%	0%	0%	0%	0%	0%
beta	0%	0%	0%	0%	0%	0%	0%	0%	0%
retvol	0%	0%	0%	0%	0%	0%	0%	0%	0%
betasq	0%	0%	0%	0%	0%	0%	0%	0%	0%
absacc	0%	0%	0%	0%	0%	0%	0%	0%	0%
acc	0%	0%	0%	0%	0%	0%	0%	0%	0%

	2015	2016	2017	2018	2019	2020	2021	All Years
securedind	100%	100%	100%	100%	100%	100%	96%	99%
mom1m	0%	0%	0%	0%	0%	0%	0%	1%
convind	0%	0%	0%	0%	0%	0%	4%	0%
mom12m	0%	0%	0%	0%	0%	0%	0%	0%
lev	0%	0%	0%	0%	0%	0%	1%	0%
beta	0%	0%	0%	0%	0%	0%	0%	0%
retvol	0%	0%	0%	0%	0%	0%	0%	0%
betasq	0%	0%	0%	0%	0%	0%	0%	0%
absacc	0%	0%	0%	0%	0%	0%	0%	0%
acc	0%	0%	0%	0%	0%	0%	0%	0%

Notes: This table reports the top 10 most influential variables for the ENet model, by year.

Table H.3: Top 10 relative variable importance - ERT

	2006	2007	2008	2009	2010	2011	2012	2013	2014
securedind	28%	28%	27%	50%	45%	45%	46%	46%	46%
betasq	4%	4%	2%	4%	4%	4%	5%	5%	5%
mom12m	3%	3%	4%	4%	5%	4%	4%	4%	4%
beta	2%	2%	1%	3%	3%	4%	3%	3%	3%
baspread	3%	3%	3%	3%	3%	3%	3%	3%	3%
mom1m	12%	11%	7%	0%	1%	2%	2%	2%	2%
maxret	3%	3%	2%	2%	2%	3%	3%	3%	3%
retvol	2%	2%	3%	2%	2%	2%	2%	2%	2%
dy	1%	1%	2%	1%	1%	1%	1%	1%	1%
nincr	1%	1%	0%	2%	2%	2%	2%	2%	2%

	2015	2016	2017	2018	2019	2020	2021	All Years
securedind	46%	46%	46%	45%	46%	46%	40%	43%
betasq	5%	5%	5%	5%	5%	5%	3%	4%
mom12m	4%	4%	4%	4%	4%	4%	3%	4%
beta	3%	3%	3%	3%	3%	3%	3%	3%
baspread	3%	3%	3%	3%	3%	3%	3%	3%
mom1m	2%	2%	2%	2%	2%	2%	1%	2%
maxret	3%	3%	3%	3%	3%	3%	1%	2%
retvol	2%	2%	2%	2%	2%	2%	1%	2%
dy	1%	1%	1%	1%	2%	2%	7%	2%
nincr	2%	2%	2%	2%	2%	2%	2%	2%

Notes: This table reports the top 10 most influential variables for the ERT model, by year.

Table H.4: Top 10 relative variable importance - NN1

	2006	2007	2008	2009	2010	2011	2012	2013	2014
securedind	29%	22%	45%	102%	69%	53%	7%	59%	65%
mom1m	19%	29%	20%	26%	20%	16%	17%	14%	1%
betasq	0%	0%	0%	0%	0%	0%	1%	0%	2%
mom12m	9%	10%	6%	8%	5%	4%	16%	6%	3%
beta	0%	0%	0%	0%	0%	0%	1%	0%	1%
chmom	8%	10%	3%	9%	6%	7%	1%	2%	1%
dy	2%	2%	2%	5%	2%	4%	6%	5%	3%
indmom	0%	0%	0%	0%	0%	0%	0%	0%	3%
mom6m	2%	4%	2%	5%	4%	5%	13%	2%	2%
retvol	3%	1%	1%	-5%	1%	1%	3%	1%	1%

	2015	2016	2017	2018	2019	2020	2021	All Years
securedind	80%	79%	84%	92%	83%	91%	91%	81%
mom1m	3%	3%	3%	3%	3%	4%	4%	4%
betasq	2%	3%	2%	2%	3%	2%	1%	2%
mom12m	2%	2%	1%	1%	1%	1%	1%	2%
beta	2%	2%	2%	3%	2%	2%	2%	2%
chmom	1%	1%	1%	2%	2%	2%	2%	2%
dy	2%	2%	1%	1%	1%	0%	1%	1%
indmom	2%	1%	1%	1%	2%	1%	1%	1%
mom6m	1%	0%	1%	1%	1%	1%	1%	1%
retvol	1%	1%	2%	1%	1%	1%	1%	1%

Notes: This table reports the top 10 most influential variables for the NN1 model, by year.

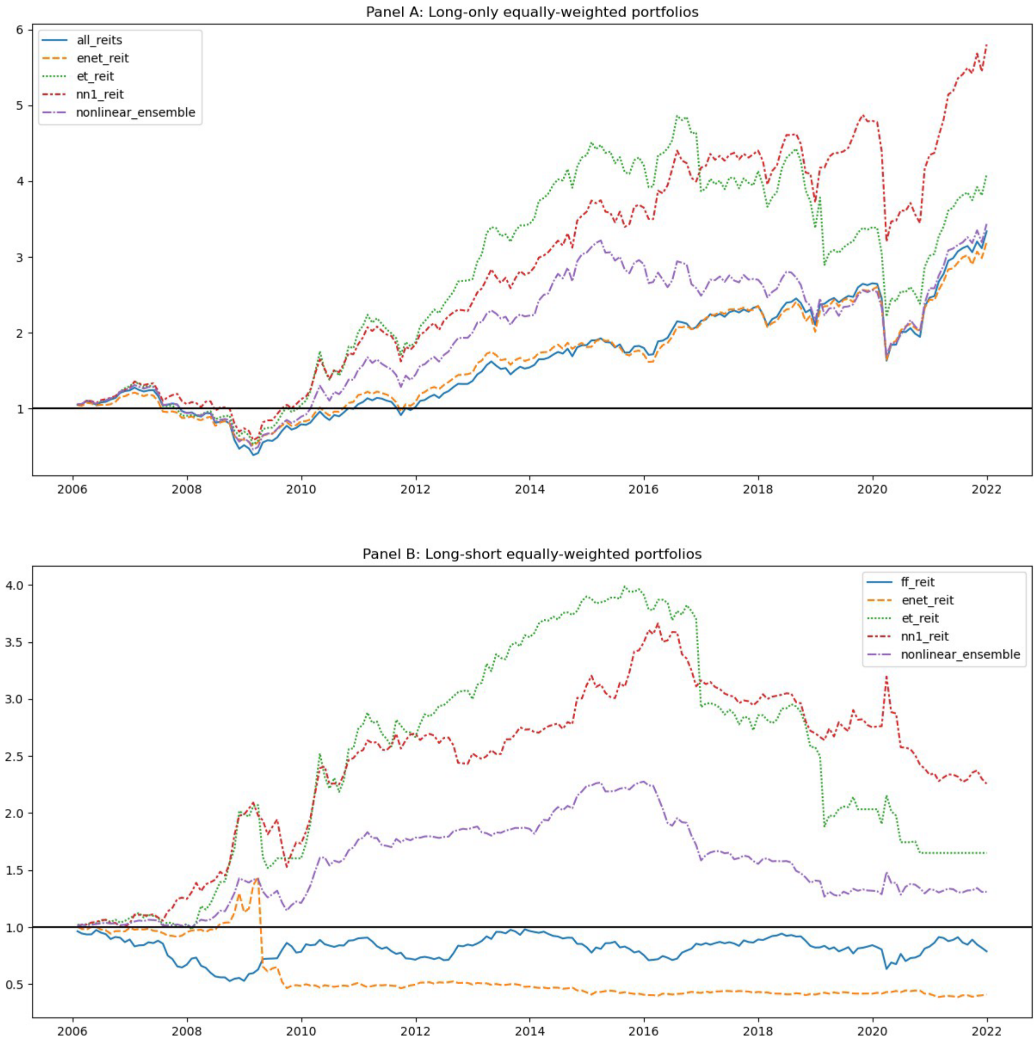
I Performance of equally-weighted machine learning portfolios

Table I.1: Performance of equally-weighted machine learning portfolios

	Avg	SD	S.R.	t-stat	Skew.	Kurt.	Max DD	Max 1M Loss	Corr
Panel A: Long-only, equally-weighted portfolio									
All REITs	0.85	6.45	0.45	6.30	-0.67	9.36	-69.85	-33.58	1.00
ENet	0.76	5.40	0.49	6.75	-1.26	7.01	-58.09	-30.75	0.93
ERT	0.98	6.92	0.49	6.80	-0.44	3.54	-60.87	-27.43	0.87
NN1	1.11	6.14	0.63	8.67	-0.07	6.01	-57.14	-27.16	0.95
Nonlinear Ensemble	0.85	6.42	0.46	6.38	-0.23	4.79	-65.08	-27.42	0.94
Panel B: Long-short, equally-weighted portfolio									
OLS-2	-0.03	4.37	-0.02	-0.31	-0.17	3.29	-45.94	-21.29	0.31
ENet	-0.28	5.19	-0.19	-2.58	-6.06	64.29	-73.21	-54.59	-0.56
ERT	0.39	4.94	0.27	3.76	-0.74	8.27	-58.59	-24.64	-0.28
NN1	0.50	3.79	0.45	6.27	0.55	3.30	-38.40	-11.86	-0.31
Nonlinear Ensemble	0.19	3.10	0.21	2.91	0.46	3.83	-44.32	-9.66	-0.34

Notes: This table reports the out-of-sample performance measures for the best performing machine learning models of the equally-weighted long-only and long-short portfolios based on the full sample. "Avg" : average realized monthly return (%). "Std": the standard deviation of realized monthly returns (%). "S.R.": annualized Sharpe ratio. "T-stat": t-statistic of realized monthly returns. "Skew": skewness. "Kurt": kurtosis. "MaxDD":the portfolio maximum drawdown (%). "Max 1M Loss": the most extreme negative realized monthly return (%). "Corr": correlation of realized monthly returns against the All REITs benchmark returns. In Panel A, the portfolios are based on a long-only strategy of holding REITs with the highest expected returns (top 30 percent), and the benchmark portfolio is the weighted index of all REITs in the sample period. In Panel B, the portfolios are based on a long-short strategy of buying REITs with the highest expected returns (top 30 percent) and shorting REITs with the lowest expected returns (bottom 30 percent), and the benchmark is a long-short portfolio based on predicted returns from OLS-2. Nonlinear ensemble refers to a grand ensemble of all nonlinear methods in our machine learning toolkit, comprising of RF, GBRT, ERT, NN1, NN2, NN3, NN4, and NN5. All portfolios are equally-weighted.

Figure I.1: Cumulative return of machine learning portfolios (equally weighted)



Notes: This figure shows the cumulative returns of the best performing machine learning portfolios. In Panel A, the portfolios are based on a long-only strategy of holding REITs in the top quintile, and the benchmark portfolio is the weighted index of all REITs in the sample period. In Panel B, the portfolios are based on a long- short strategy of buying the highest expected return REITS (top quintile) and shorting the lowest expected return REITs (bottom quintile), and the benchmark is a long-short portfolio based on predicted returns from OLS-2. Nonlinear_ensemble refers to a grand ensemble of all nonlinear methods in our machine learning toolkit, comprising RF, GBRT, ERT, NN1, NN2, NN3, NN4, and NN5. All portfolios are equally-weighted.