CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Economic Studies

Mispricing in leveraged value small-capitalization stocks

Master's thesis

Author: Bc. Jan Picálek

Study program: Economics and Finance

Supervisor: Mgr. Martin Hronec

Year of defense: 2022

Declaration of Authorship

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, May 2, 2022

Bc. Jan Picálek

Abstract

We study returns in the universe of leveraged value small-capitalization stocks, a universe with historically significant exposure to common risk factors. We separate future winners and losers within this universe of risky stocks by adopting machine-learning-based mispricing strategy. The strategy considers 34 stock-level characteristics to predict 1-month-ahead returns and construct a long-short portfolio accordingly. The portfolio yields abnormal risk-adjusted returns of 0.42% per month out-of-sample, uncovering statistically significant mispricing. The machine-learning algorithm is trained on leveraged value small-capitalization stocks, so it captures universe-specific nonlinearities and variable interactions. The nonlinear effects and predictive power of individual variables are extracted and presented as well. We found no evidence of a relationship between the magnitude of the mispricing and credit cycles, or market volatility.

JEL Classification	G11, G12, G14,
Keywords	Anomalies, Predictability of returns, Asset pricing tests, Leveraged equities, Value stocks
Title	Mispricing in leveraged value small-capitalization stocks

Abstrakt

Zkoumáme výnosy hodnotových akcií s vysokým zadlužením a nízkou tržní kapitalizací, tj. akcií s historicky značnou expozicí vůči běžným rizikovým faktorům. Za použití strojového učení vybíráme z množiny těchto rizikových akcií ty, jež by se měly v budoucnu nadměrně zhodnotit. V rámci této strategie zohledňujeme 34 akciových charakteristik a predikujeme budoucí výnosy jednotlivých akcií, na jejichž základě pak každý měsíc sestavujeme long-short portfolio. Nadměrná výnosnost strategie 0.42% za měsíc i přes úpravu o riziko na testovacím vzorku dat ukazuje, že chybné ocenění je statisticky signifikantní. Použitý algoritmus strojového učení se učil na množině hodnotových akciích s vysokým dluhem a nízkou tržní kapitalizací, a zachycuje vztahy specifické pro tuto množinu, včetně vztahů nelineárních a interakcí jednotlivých proměnných. Tyto nelineární vztahy a prediktivní schopnost jednotlivých proměnných jsou extrahovány a následně i prezentovány. Mezi mírou chybného ocenění a kreditními cykly či tržní volatilitou jsme nenašli žádnou spojitost.

Klasifikace JEL	G11, G12, G14,
Klíčová slova	Anomálie, Prediktabilita výnosů, Testy oceňovacích modelů, Zadlužené společnosti, Hodnotové akcie
Název práce	Chybné ocenění akcií s nízkou tržní kapitalizací a vysokým dluhem

Acknowledgments

The author would like to express his deepest gratitude to his supervisor Mgr. Martin Hronec for his invaluable guidance provided throughout the entire thesis process.

Bibliographic Record

Picálek, Jan: *Mispricing in leveraged value small-capitalization stocks.* Master's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2022, pages 83. Advisor: Mgr. Martin Hronec

Contents

List of Tables

List of Figures

Acronyms

CAPM Capital asset pricing model

CML Capital Market Line

NYSE New York Stock Exchange

AMEX American Stock Exchange

Nasdaq National Association of Securities Dealers Automated Quotations

COGS Cost of goods sold

CAPEX Capital expenditures

IPO Initial public offering

SEO Seasoned equity offering

IFRS International Financial Reporting Standards

US GAAP United States Generally Accepted Accounting Principles

OLS Ordinary Least Squares

NAICS North American Industry Classification System

LTM Last twelve months

NTM Next twelve months

FRED Federal Reserve Economic Data

GBM Gradient Boosting Machines

Master's Thesis Proposal

Author	Bc. Jan Picálek
Supervisor	Mgr. Martin Hronec
Proposed topic	Mispricing in leveraged value small-capitalization stocks

Motivation

Asset prices should equal the sum of discounted future cash flows. Discounted, so when valuing an asset, the pitfall is not only to correctly predict the future cash flows but to determine the exact discount rate such that investors level of risk aversion and opportunity cost are really reflected. Capital asset pricing model (CAPM) by Sharpe (1964) and Lintner (1965) had served for this purpose for decades. However nowadays, there is substantial amount of evidence on various factors that capture more variance than the single variable - market proxy of CAPM. Examples of the most prominent factors affecting returns on top of the CAPM framework include size effect noted by Banz (1981), leverage effect identified by Bhandari (1988) or value effect uncovered by Chan, Hamao & Lakonishok (1991) and many other anomalies. As a result of the ability of those factors to further explain returns, multi-factor models taking the implied premiums into account are preferred in asset pricing literature. The mostly referred to are three-factor or five-factor models by Fama & French (1992) and (2015) respectively. In spite of the general acceptation of these models, plenty of anomalies is still left behind and non-zero α-returns can be observed even in these models if the right factors and strategies are applied, e.g. reflecting intangibles along with fundamental analysis yields positive α even in five-factor model with augmented momentum factor (Eisfeldt et al., 2020). On the other hand, pool of academic literature focused on detecting anomalies and successful trading strategies has became very saturated and more importantly the historically discovered patterns are getting reflected in the asset prices (Mclean Pontiff, 2016). Example of such factor whose premium vanished over the course of last decade is the "value effect".

I would like to take a closer look at companies exposed to some of the historically most impactful factors mentioned above - value, size and leverage, and investigate whether 1) such universe still "enjoys" the premiums it used to and more importantly 2) there are at least some "unexplored betas" within the universe, e.g. if we are able to consistently pick firms with excess risk-adjusted returns based on various factors, there must be a variation not captured by standard asset pricing models such as five-factor model etc. The latter is the focal point of this thesis as there is a lot of space for potential mispricing by investors in this universe.

Since Piotroski (2000) argues that success of the value strategy, portfolio with high book-to-market ratio in this case, originates from the high return of only a few companies while majority does not even produce positive risk-adjusted returns, some of the value companies are probably heavily mispriced. In case of leveraged firms, it is their ability to repay their debt that separates the winners from losers (Chingono & Rasmussen, 2016) and this ability might be subject to incorrect assessment by investors, especially with respect to different credit cycles. In addition, firms with smaller market capitalization are subject to scarce coverage by sell-side analysts increasing the probability of mispricing by retail investors by incorrect prediction of future cash flows.

Leveraged firms with favourable valuation multiples and small market capitalization appear to create the pool of stocks 1) whose discount rate truly reflecting investors preferences might by subject to yet undocumented factors and 2) where incorrect evaluation of future cash flows by investors might occur. Therefore, I would like to construct stock ranking system within such pool of stocks with emphasis on ability to repay debt and fundamental analysis to show if this is the right environment to implement mispricing strategy.

Hypotheses

> Hypothesis #1: It is possible to predict firm's ability to pay down debt in the universe of leveraged small-capitalization firms.
>
> Hypothesis #2: Investing strategy based on the constructed stocks ranking mechanism yields excess risk-adjusted returns.
>
> Hypothesis #3: Mispricing by investors regarding the leveraged small-cap high value firms differs with credit cycles. E.g. the constructed stocks picking system performs better during periods of high credit spread.

Methodology

Yearly or quarterly international cross-section data from the most recent decades from selected markets will be used for the purpose of this study. Both company level and macroeconomic data will be extracted from Thomson Reuters Datastream.

First step in the analysis is to select the criteria for the universe of leveraged value small-caps. Even though small-capitalization firms are usually defined as those with total equity value between USD 300 milion and USD 2 billion, the definition in this study might slightly differ such that the resulting number of firms in the constructed universe is appropriate for the analysis. Subsequently, the companies will be ordered based on a debt-related ratio and a valuation multiple in order to determine those leveraged and containing value, most likely based on a percentile boundary. As a result of the intersection of the small, value and leveraged sets of companies, universe of stocks for further analysis will be created.

Since debt reduction is the most important factor in the process detecting mispriced companies within the leveraged value small-caps (Chingono & Rasmussen, 2015), separate model for prediction of the probability of debt repayment is to be constructed. Panel data linear regression and machine-learning techniques will be applied to company financial data and forward-looking estimates of analysts. The following equation will be estimated:

$$y_{i,t} = f(x_{i,t-1,1}, x_{i,t-1,2}, .., x_{i,t-1,K}) + \epsilon_{i,t} \tag{1}$$

where y is dummy variable equal to 1 if the company i reduces its long-term debt in period t compared to previous period and 0 otherwise. x_k denotes individual explanatory lagged variables, e.g. gross margin, growth rate of sales, asset turnover etc.

As the intended stocks ranking mechanism aims to predict and not to explain, again machine-learning techniques such as random forests and gradient boosting machines will be applied to predict returns in the subsequent period. It is defined as follows:

$$r_{i,t} = g(y_{i,t}, \mathrm{c}_{t-1}, x'_{i,t-1,1}, x'_{i,t-1,2}, .., x'_{i,t-1,K}) + u_{i,t} \tag{2}$$

where $r_{i,t}$ is the return on stock i during period t, c_{t-1} is a variable capturing the efficiency of credit market in period t-1, $y_{i,t}$ is the estimated probability from equation *(1)* and x'_k denotes individual explanatory lagged variables.

Such and estimation procedure will be applied on T periods representing the in-sample part of dataset.

In the final stage, portfolio consisting of certain amount of stocks with the highest predicted returns based on equation *(2)* will be constructed and rebalanced each year in the out-of-sample part of dataset.

Performance testing of the constructed ranking system will be based on CAPM and multifactor models by Fama & French (1992,2005) using the entire out-of-sample period.

Expected Contribution

Historically we have been flooded with factors explaining the cross-sectional variation in returns. However, in many cases the findings can be attributed to specific datasets or the implied premiums just vanished through the following years as investors risk aversion and opportunity cost are functions and not constants. For example, the near-zero interest rate environment we have observed in recent years might have forced them to explore other investing options. Therefore, I would like to contribute to the existing asset pricing research with machine-learning stocks ranking system based on prediction of discount rates in such universe of stocks that have historically been positioned to yield premiums but now only rarely are found. Especially, the introduction of credit cycle explanatory variable in the prediction might unveil some variation in returns of value and leveraged equities. Moreover, I would like to reflect such relationships that might be only universe specific and identify companies that are most suitable for mispricing strategy as investors tend to incorrectly predict their future performance.

Analogous study focused on similar universe of stocks is Chingono & Rasmussen (2015, 2016). However, substantial differences and extensions should be noted. Especially, our machine-learning approach for construction of the stocks ranking system is expected to perform better out-of-sample due to the predictive power of this technique. Moreover, this study extends the geographical coverage as we are going to use international data so differences in predictability and repayment risks between individual markets could be observed. There will be also update and extension in terms of data recency so more up to date relationships could be taken into account during estimation and out-of-sample testing. Last but not least introduction of credit cycles in the models is expected offer better insight to investors risk aversion variability in time resulting in more accurate discount rates.

Outline

1. Introduction
2. Asset pricing theory and anomalies review
3. Data & Universe construction
4. Methodology
5. Results
6. Conclusion

Core bibliography

Markowitz, H. (1952). Portfolio Selection. The Journal of Finance, 7(1), 77-91.

Lintner, J. (1965). The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. The Review of Economics and Statistics, 47(1), 13-37.

Sharpe, W.F. (1964). Capital asset prices: a theory of market equilibrium under conditions of risk. The Journal of Finance, 19: 425-442.

Chingono B, Rasmussen D. (2015). Leveraged small value equities.

Chan, L.K.C., Hamao, Y. and Lakonishok, J. (1991). Fundamentals and Stock Returns in Japan. The Journal of Finance, 46: 1739-1764

Fama, E.F. and French, K.R. (1992). The Cross-Section of Expected Stock Returns. The Journal of Finance, 47: 427-465.

Chingono B, Rasmussen D.(2016). Forecasting debt paydown among leveraged equities.

Piotroski, J. (2000). Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers. Journal of Accounting Research, 38, 1-41.

Bhandari, L.C. (1988). Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence. The Journal of Finance, 43: 507-528.

Fama, E.F. and French K.R. (2015). A five-factor asset pricing model. Journal of Financial Economics, 116(1).

Banz, Rolf W. (1981). The relationship between return and market value of common stocks, Journal of Financial Economics, 9, issue 1, p. 3-18.

Eisfeldt, A. L., Kim, E., Papanikolaou, D. (2020). Intangible value. NBER Working Paper Series, (28056).

Mclean, R. D., Pontiff, J. (2016). Does Academic Research Destroy Stock Return Predictability? Journal of Finance, 71(1), 5–32.

1 Introduction

Asset prices should equal sum of discounted expected future cash flows. Determination of the rate used for the discounting has become the leading discipline in the asset pricing literature since the introduction of Capital asset pricing model (CAPM) by Sharpe (1964) and Lintner (1965). The assumption that risk in the cross-section is entirely captured by beta on market risk premium has been successfully challenged ever since, leading to emergence of factor models that contain additional risk proxies such as value, size, or profitability. Still, there has been a flooding of new anomalies[1] that yielded abnormal returns even after accounting for the generally recognized risk factors, examples of such anomalies include intangibles, momentum effect, past losers effect, investment policy. Hou *et al.* (2020) counts up to 450 anomalies identified by prior literature, however, majority of them cannot be replicated anymore. This suggests that nowadays, the market is more efficient and more of the publicly available information is getting incorporated into stock prices. Post-publication decline in anomaly alpha returns further supports the assertion that investors quickly trade upon new information (Mclean & Pontiff, 2016). This suggests that nowadays, to uncover mispricing, traditional approach of sorting stocks on a single anomaly is not enough anymore. Jacobs & Müller (2018) points out to dimensionality of stock returns and show that aggregation of more than 200 anomalies into a single predictive signal is superior to the traditional one-dimensional approach. In addition to that, recent anomalies research that combined the anomalies with machine-learning techniques further outperformed out-of-sample substantially. (see Gu *et al.*, 2020; Tobek & Hronec, 2021). Vast majority of the literature is based on as wide spectrum of stocks as possible, leaving potential subcategory-specific relationships of stocks undiscovered. For example, deleveraging is the main driver of abnormal risk-

[1] In this thesis, term factor primarily denotes a risk-related determinant of expected return while anomaly refers to a rather behavioral-based effect. Though, in many cases, the terms factor, anomaly, signal, and stock characteristic are used interchangeably.

adjusted returns among leveraged value small-capitalization stocks (Rasmussen & Chingono, 2015).

We build upon the evidence that stock returns are dimensional (Jacobs & Müller, 2018), and that machine-learning provides additional predictive power compared to OLS-based regression methodology (Gu *et al.*, 2020). Using gradient boosting machines algorithm, we consider 34 lagged stock-level characteristics to uncover mispricing among leveraged value small-capitalization stocks. The critical element of this thesis is the focus on the universe of leveraged value small-capitalization stocks, which is underserved by prior research on the cross-section of expected returns, especially when combined with machine-learning-based methods. Lower reporting quality, lack of sell-side research, leverage, and higher probability of financial distress make this risky universe a good candidate for a naive assessment by investors. Piotroski (2000) arguing that only a few positive outliers drive the success of the value strategy is an another call for further separation of the future winners and losers among such stocks. Application of the machine-learning-based methodology allows us to reflect more complex and conditional relationships that might be present in this particular universe due to its exposure to the prominent risk factors and mispricing potential. In addition to reflecting such relationships in the prediction task, we extract these often non-linear effects for the most important predictors and present them to the reader. We also further expand upon the mispricing strategy by further investigating whether the magnitude of the mispricing is influenced by prevailing credit conditions or market volatility. The entire analysis is based on relatively recent data (i.e. since 2000) to reflect primarily contemporary behaviour of investors.

Our strategy yields out-of-sample abnormal return of about 0.42% (0.34 %) per month after adjusting for three (five) risk factors of Fama & French, i.e. we uncover mispricing in the universe of leveraged value small-capitalization stocks. Majority of the abnormal returns of the mispricing-based long-short portfolio are driven by the short leg, so the algorithm is more successful in detecting the overvalued stocks. We also find that behavioral variables (such as return in the past six months) are among those with greatest predictive power, and that the effects of individual variables are usually nonlinear. We find no evidence of a relationship between the magnitude of the mispricing and credit cycles, or market volatility. We also predicted the ability to deleverage in the next twelve months with out-of-sample accuracy of 62%. As opposed to Rasmussen & Chingono (2015), the future deleveraging did not help much to explain future stock returns.

The remainder of this thesis is organized as follows. Section 2 reviews key asset pricing concepts and discusses the predictability of stock returns in the cross-section. Dataset description and construction of the universe are provided in Section 3. Section 4 summarizes the research design and defines the methods employed within individual steps of the analysis. Section 5 presents empirical findings and Section 6 concludes.

2 Asset Pricing Theory and Anomalies Review

2.1 Fundamental Asset Pricing Concepts

Markowitz (1952) formed the crucial theory of portfolio selection that serves as a building block for asset pricing concepts that have been widely accepted and used for decades, especially CAPM introduced by Sharpe (1964) and Lintner (1965). Markowitz analytically demonstrated in his seminal work that portfolio diversification can reduce risk without any harm on expected return as a result of imperfect correlations between assets. It shows that expected returns in a portfolio are combined linearly while risk of the portfolio is not. It is nonlinear since the asset returns do not move perfectly along each other. Markowitz also introduced graphical illustration of this theory called "efficient frontier", depicting all the risky portfolios yielding maximum expected return for a given level risk or *vice versa*.

Under assumption of buying and borrowing of risky assets as well as lending and borrowing at risk-free rate, there is only one optimal portfolio consisting of solely risky assets for risk-averse rational investors - the market portfolio (Tobin, 1958). This efficient portfolio is the one with the most favourable reward-risk relationship, usually referred to as Sharpe ratio[1] (Sharpe, 1966). Tobin separated the problem of finding the optimal risky portfolio in two phases: 1) finding the optimal portfolio of solely risky securities lying on efficient frontier as developed by Markowitz (1952) and 2) determining the optimal combination of the risky portfolio constructed in step *1)* and risk-free asset. In practice positions in risk-free asset are taken either by buying government bonds (long

[1]Sharpe ratio $= \frac{E(r_i)-r_f}{\sigma_i}$
where $E(r_i)$ is the expected return on asset i, r_f is risk-free rate and σ_i is standard deviation of asset i.

case) or by buying the optimal risky portfolio on margin (short case) and they serve to decrease or increase risk, respectively. As a result, rational investors are supposed to hold a combination of market portfolio (as it is the one exposed only to systematic risk due to diversification) and a position in risk-free asset (to achieve optimal risk profile determined by degree of risk-aversion of individual investor). Expected returns and standard deviations of all these potential combinations are captured by Capital Market Line (CML).

Building on top of the conclusions above regarding portfolio selection on the level of individual investor, Sharpe (1964) derived capital market equilibrium. Under the assumption of rationality of investors and homogeneity of their expectations, all investors will invest in one particular combination of risky securities. In response to the high demand for such assets and no demand for alternative investments, their prices will rise and fall, respectively. Such change in asset prices will naturally affect expected returns leading to subsequent change in the composition of optimal risky portfolio. This interplay will continue until all assets are cleared - purchased by someone. This occurs when all asset are priced such that they are included in at least one[2] efficient portfolio. Then, equilibrium of capital market is reached and no efficient risky portfolio will be rebalanced any further. Therefore, there is a linear relationship between risk and expected return for efficient risky portfolios and these will be composed of only risky assets or risky assets combined with risk-free asset. In either case, all these portfolios will be perfectly positively correlated (Sharpe, 1964).

After determining the process of convergence to equilibrium asset prices, Sharpe (1964) moved on deriving individual asset returns relative to an efficient portfolio (the given individual asset must be part of). As the efficient portfolios are fully diversified and exposed only to general risk of the entire economy, the relationship between returns of efficient portfolios and returns of an individual asset represents its responsiveness to changes in economic activity. The other type of risk of individual assets, *firm-specific risk*, can be always eliminated by diversification. For that reason, it is irrelevant to take this type of risk into account when determining expected returns of an asset. Moreover, the relationship between the responsiveness to economic activity proxied by efficient

[2]Sharpe (1964) argued that there can be multiple efficient risky portfolios as long as they are perfectly positively correlated, as opposed to Tobin's conclusion of only single efficient risky portfolio.

portfolio rate of return and an individual asset rate of return will be linear in equilibrium as any non-linearity would be exploited by investors (Sharpe, 1964).

Summarizing the contributions to CAPM made by several authors, the model is generally defined in literature as follows:

$$E(r_i) = r_f + \beta_i(r_m - r_f) \tag{2.1}$$

where r_i is the rate of return on asset i, r_f is the risk-free rate, r_m is the rate of return of efficient risky portfolio, the market portfolio, and β_i represents responsiveness of returns on asset i to overall economic activity represented by excess returns on market portfolio and it is defined as

$$\beta_i = \frac{Cov(r_i, r_m - r_f)}{Var(r_m - r_f)} \tag{2.2}$$

There are several commonly used assumptions for equation (2.1) to hold. First, investors are rational, reluctant to risk and select investments maximizing their expected returns for their preferred level of risk. Second, investors' expectations regarding asset returns, standard deviations and mutual correlations are homogeneous. Third, capital markets exhibit certain characteristics such as no transaction costs, no taxes, no short-selling restrictions and existence of risk-free assets. Moreover, the capital markets are competitive and non-discriminating, i.e. all the characteristics apply to every market participant and every investor is a price taker. Fourth, any investor may invest to or initiate a short position of any size, even fractional position, of any asset, including the risk-free asset.

Given the assumption requirements for formula (2.1), its applicability in real world seems questionable. It may be the "no taxes" assumption that raises the doubts of investors, "no transaction costs" or any other. For example, tax-loss harvesting or holding successful positions for an extended period to avoid capital gain tax explain why investor behaviour is affected by taxes (Constantinides, 1983). Similarly, other real-world examples violating the other CAPM assumptions could be found. Therefore, there is considerable amount of literature focused on testing of CAPM at various markets or under relaxation of certain assumptions.

One of the early tests of the model was conducted by Jensen *et al.* (1972) on a sample of all stocks listed on NYSE between 1926 and 1966 when they applied both cross-sectional and time-series regressions. In their time-series regression they estimated the following equation:

$$R_{j,t} = \alpha + \beta_j R_{mt} + \epsilon_{j,t} \tag{2.3}$$

Since they worked with excess returns instead of total returns, significantly non-zero intercept, α, would imply rejection of CAPM. To cope with the issue of cross-sectionally dependent errors $\epsilon_{j,t}$ in case of individual securities data, they aggregated securities into portfolios ranked by their respective βs estimated from period preceding the main regression period. So j in equation (2.3) denotes portfolios constructed based on past relative correlation of the included securities and the market. Even though their results were rather mixed, negative relationship between estimated α and portfolio βs was apparent. They showed that more risky securities yielded less than implied by theory and *vice versa*. In case of cross-sectional regression using the same grouping procedure, they concluded that relation between average excess returns and β is linear, however, the results were inconsistent with theory due to variation through different estimation periods. Based on the above mentioned findings, Jensen *et al.* (1972) claimed that CAPM theory is not substantiated by real data.

Another early test, test by Fama & MacBeth (1973) aimed on testing the implications of CAPM. Namely, linearity between risk and return, β being the ultimate measure of risk capturing all systematic risk and positivity of the trade-off between risk and return. The following equation on β-ranked portfolios was estimated:

$$R_{pt} = \hat{\gamma}_{0t} + \hat{\gamma}_{1t}\hat{\beta}_{p,t-1} + \hat{\gamma}_{2t}\hat{\beta}^2_{p,t-1} + \hat{\gamma}_{3t}\bar{s}_{p,t-1}(\hat{\epsilon}_i) + \hat{\eta}_{p,t} \tag{2.4}$$

where $\hat{\gamma}_{2t} = 0$ hypothesized linearity of risk-return relation, $\hat{\gamma}_{3,t} = 0$ hypothesized uniqueness of β as a measure of market risk and $\hat{\gamma}_{1t} > 0$ hypothesized that investors are rewarded for bearing additional systematic risk. Their results came in favor of classical market model and its implications, concluding that investors hold efficient asset combinations in terms of risk-return trade-off.

In spite of very restrictive assumptions that are not deemed realistic and mixed performance in terms of empirical testing, CAPM is due to its simplicity even nowadays still widely used in practice, especially in discount rate determination for the purposes of valuation and capital budgeting.

2.2 Multi-factor Models

Despite its strong theoretical background built on Markowitz's portfolio choice theory, CAPM's ability to explain variation in cross-section of rate of returns was in question since its formulation by Sharpe (1964), Lintner (1965). Subsequent research centered especially around the fact that β was supposed to represent all the risk associated with owning a stock. In the wake of this research, other factors capturing the cross-sectional variation of stock returns along with β started to emerge.

Among the first factors capturing cross-sectional variation of returns on top of the single-variable β specification was size, usually defined as firm's market capitalization, documented by Banz (1981), Keim (1983), Fama & French (1992; 1993) and many others. The rationale behind the size premium provided by Fama & French (1992) and consistent with Chan & Chen (1991) is that small companies are more prone to financial distress and investors demand to be compensated for such an extra risk. This view regarding the source of the size premium is coherent with the fact that size effect is strongest in case of firms with high probability of default (Vassalou & Xing, 2004). Another potential explanation is found in a greater exposure to information asymmetry as reporting of smaller firms is less thorough (Banz, 1981) or lower liquidity of small-capitalization stocks (Amihud & Mendelson, 1986). There is plenty of research discrediting the idea of size serving as a risk proxy, e.g. size effect being caused primarily by outliers (Knez & Ready, 1997) or increased magnitude of the size effect in January (Keim, 1983) etc.

Value is another factor well documented to explain cross-sectional variation in returns. However, the source of this relation remains a subject of heated debate. Proxied by various variables in prior literature, the most prominent one with consistent and robust effect is book-to-market ratio (B/M, defined as book value of equity divided by market capitalization) found in several studies including Chan *et al.* (1991), Rosenberg *et al.* (1985), Fama & French (1992; 1993), Lakonishok *et al.* (1994) etc. The ability of B/M to further explain returns is robust to inclusion of other value factors in a regression, i.e. B/M effect is not consumed by the effects of the other value-related factors. The list of other variables "mimicking" value effect and explaining variation in cross-section of returns on top of CAPM extends primarily to cashflow-to-price ratio (Chan *et al.*, 1991; Davis, 1994) and earnings-to-price ratio (E/P, net earnings divided by market capitalization) (Basu, 1983; Jaffe *et al.*, 1989; Cook & Rozeff, 1984; Reinganum, 1981; Davis, 1994). Earnings-to-price ratio, however, yields diverse results across various studies. Whereas Basu (1983) claims that size effect disappears when added as independent variable along the E/P, Reinganum (1981) maintains the opposite stance that it is the E/P effect that is not robust to size.

Disagreement whether the relationships of returns and certain factors are present due to ability of such factors to proxy for risk or whether these effects are rather behavioural still persists. Fama & French (1992) argue that risk associated with equities has multiple dimensions and that they could be captured by market capitalization and B/M. Value, proxied by B/M, reflects investors' expectations of future performance and probability of potential distress translating into higher rate of return required by equity investors as a compensation for the additional risk, i.e. higher B/M implies higher expected returns as those companies have weak future outlook and are exposed to increased risk (Fama & French, 1992). Alternatively, Lakonishok *et al.* (1994) shows that value effect is caused by naive investors extrapolating weak past growth too far into the future. They conclude that value investing does not come with an additional risk to be compensated for as the value stocks do not exhibit any significantly higher standard deviations or βs compared to growth stocks. The naive investors are, on the other hand, too optimistic about future prospects of past well-performers making the growth stocks overpriced relatively to value stocks. This explanation of Lakonishok *et al.* (1994) is consistent with overreaction rationale suggested by Bondt & Thaler (1985). In their study, past losers outperformed past winners by 25% during a 3-year period following portfolio formation despite carrying less volatility. Haugen (1995) also sides with behavioural explanation and points out to overreaction of investors to new information as they incorrectly believe that current extraordinary performance will persist too long.

Despite contradicting claims, whether value and size effects are risk-based or behavioural, there is enough evidence of their ability to capture the cross-sectional variation of returns.

Fama & French (1992) tested several risk factors identified in prior literature B/M, E/P, size, β and leverage. Except β, all of the tested variables exhibit explanatory power when studied individually. Including all of them simultaneously, B/M and size apparently consumed the effects of E/P and leverage. These findings served as a backbone for their later study, Fama & French (1993), to extend CAPM for value and size effects - three-factor model.

In Fama & French (1993), the methodology is distinct to the regressions they used in Fama & MacBeth (1973) and the one in Fama & French (1992). First, they sort their sample of stocks listed on NYSE, AMEX and Nasdaq between 1963-1991 into two groups based on size and three groups based on B/M ratio. Based on the intersections of these five groups, six (2x3 sorts) portfolios are constructed, i.e. one portfolio is composed of stocks with low market capitalization and high M/B, another consists of stocks with high market capitalization and low M/B etc. Subsequently, explanatory variables hypothesized to proxy for common risk factors, SMB and HML are created. Each month SMB variable is calculated by subtracting average returns on the large market capitalization portfolios from the average returns on the small market capitalization portfolios while controlling for value effect. HML is calculated analogously for value. Formulating the three-factor model, Fama & French used time-series regression to estimate the following equation for 25 portfolios (5x5 sorts based on value and size):

$$r_{pt} = \alpha + \beta_1 r_{Mt} + \beta_2 SMB_t + \beta_3 HML_t + e_t \tag{2.5}$$

where r_m represents market excess returns over risk-free rate and r_p is portfolio excess return in month t.

As a result, the three-factor model captured over 90% of variation in returns for vast majority of the 25 portfolios regressed separately, surpassing the explanatory power of CAPM. Moreover, the regression slopes for SMB and HML exhibited the hypothesized relationships and were significant for almost all portfolios. When market premium kept as the only explanatory variable, regression intercept, α, was highly significant for high B/M portfolios while consistently increasing with value. In case of size, rather negative relationship was observed.

When SMB and HML were added to the regression, for all the regressed portfolios but 3 out of the 25 tested, the significant α-returns vanished. However, joint F-test still, although closely, rejected the zero-intercept hypothesis. On the other hand, in all but 1 portfolio, the α was not substantially different from 0 in the economic terms. This suggests that market premium, size and value jointly serve well, but not perfectly, as proxies for systematic risk (Fama & French, 1993).

Complementing to factors in three-factor model by Fama & French (1993), Novy-Marx (2013) argues that gross profitability (defined as revenues less COGS, all divided by total assets) further explains the variation in cross-section of returns. Even though high profitability is usually associated with low B/M (hence considered a growth strategy), highly profitable firms exhibit substantially increased returns. In addition, combined with value strategy, considerable improvement of those strategies could be observed in terms of both expected returns and volatility (Novy-Marx, 2013). This partially contradicts with Fama & French (1993) risk-based explanation of value factor as they consider low B/M firms to automatically exhibit low profitability. The paper also indicates that three-factor model still lacks some factors carrying a risk premium as the three-factor model test for 5 portfolios sorted on the gross profitability rejected the model in case of the lowest and highest profitable portfolios. Novy-Marx showed monotonically increasing α with gross profitability. In addition, using returns spread between the two most and least profitable portfolios exhibited significant α-returns of 0.5% per month. Haugen & Baker (1996) used lagged return on equity as a proxy for profitability in their predictive regression. Even though the model controlled for various price-related, technical, macroeconomic and growth potential-related variables; statistically and economically significant profitability-return relationship was found.

Titman *et al.* (2004) document another factor with a significant relationship with cross-section of returns - increased capital expenditures. They find that firms with increased capital expenditures tend to underperform the benchmarks. It is important to be aware of two potential biases 1) increased investment expenditures usually follow extraordinary past performance, which makes investors to irrationally misprice (Bondt & Thaler, 1985; Lakonishok *et al.*, 1994), 2) as every investment requires a financing source, there might be a positive relation between increasing CAPEX and selling equity, which in turn is associated with negative future returns (Loughran & Ritter, 1995; Bradshaw *et al.*, 2006). So Titman *et al.* conducted empirical testing while also controlling for the two negative anomalies to conclude negative *ceteris paribus* effect of increasing capital expenditures. Cooper *et al.* (2008) selects more general proxy, year-over-year growth of total assets, and documents consistent result with those of Titman *et al.* (2004). Particularly, spread of risk-adjusted returns between low and high asset growing companies amounted to 8% with high level of statistical significance. When regressed along other recognized factors such as B/M, size, mean reversal proxies, accruals etc. the asset growth was the most significant factor in predicting the future returns (Cooper *et al.*, 2008).

In the wake of prior evidence on profitability and investment factors affecting cross-section of returns described above, Fama & French (2015) introduced five-factors model, extending its predecessor - three-factor variant. The model is defined as follows:

$$R_{i,t} = \alpha_i + \beta_{1i} R_{Mt} + \beta_{2i} SMB_t + \beta_{3i} HML_t + \beta_{4i} RMW_t + \beta_{5i} CMA_t + e_{i,t} \quad (2.6)$$

where the familiar variables are defined analogously to those in equation (2.5), RMW denotes difference in monthly returns between well-diversified portfolios with strong and weak profitability and CMA denotes the difference in monthly returns between portfolios of conservatively and aggressively investing firms.

Explanatory power of the five-factor model was tested using GRS test[3]. The test rejected the hypothesis of α being equal to zero for all tested portfolios, suggesting that there are still some patterns left unexplained. However, the model performed significantly better than its predecessor, three factor variant, even in terms of other performance measures such as R-squared etc. The authors also report that the explanatory power of value factor vanished when profitability and investment proxies are included.

[3]GRS test from Gibbons *et al.* (1989) tests the hypothesis that the intercepts, α, in a time-series regression are equal to zero for all regressed portfolios.

2.3 Anomalies

We have already discussed the most prominent factors[4] affecting stock returns - β, size, value, profitability and investment. They are well-documented in prior literature and there is usually at least a discussion whether the given effect is driven by increased risk or by irrational behaviour of investors and market inefficiency. However, there is plenty of other determinants of stock returns proposed by academic research, usually referred to as anomalies.

The accounting standards IFRS and US GAAP rely on accrual method of recording revenues and expenses. Thus, revenues are recorded once the goods or services are delivered even though the cash receipt comes with certain lag. The same holds for expenses. So the accrual and cash flow components in the financial statements might vary substantially. Using solely the accrual component in financial analysis might lead to inaccurate assessment of earning power of a company, so both these components should be taken into consideration (Graham *et al.*, 1962). Sloan (1996) argues that the earnings performance, attributable to accrual part, is more likely to disappear compared to cash flow part. Sloan (1996) reports that cashflow component has significantly higher predictive power than accrual component in terms of 1-year ahead earnings. He also finds that investors are excessively attached to earnings figures and fail to incorporate complete information in financial statements into their evaluation of stock price. On the other hand, Fairfield *et al.* (2003) attributes the accrual anomaly to conservative accounting, e.g. expensing research and development costs (R&D) coupled with lower returns on new investments. Richardson *et al.* (2006) decomposed accruals into two categories - growth-related (those related to sales generation) and efficiency-related (those related to amount of assets needed to operate) showing that poor sustainability of earnings is associated primarily with increase of the efficiency-related accruals.

When determining book value, only tangible assets such as property, plant and equipment etc. are used. Intangible assets on balance sheet (e.g. trademarks, patents, copyright etc.) are usually excluded in the calculation even

[4]Even though consensus regarding source of an effect on expected returns is not always found (risk-based or behavioral), in this thesis, the effects considered risk-related are termed factors and those considered behavioral are termed anomalies. Though, in several cases the terms are used interchangeably.

though, nowadays, they represent substantial portion of the entire asset pool of modern company. One should also consider the investments into intangible assets that are not capitalized, i.e. they do not appear on balance sheet at all. These "invisible" investments usually include employee training expenses, non-capitalized R&D costs etc. As a result, firms heavily investing into R&D appear to be priced expensively relative to its non-innovative counterparts. The growing importance of intangible assets in firms' capital stock was noted by several research papers (Eisfeldt *et al.*, 2020; Belo *et al.*, 2019). These papers estimate that, nowadays, majority of corporate investments are towards intangibles. Such evidence encourages for a revision of tradition view on value factor which is usually measured by only tangible-based B/M ratio. Eisfeldt *et al.* (2020) makes such a revision by modification of HML factor of Fama & French (1993) to account for intangibles on top of the original solely tangible book value. Modified HML factor appeared to produce higher average returns while reducing volatility at the same time. Their proposed strategy of buying value firms with augmented intangibles and selling firms with only tangible value exhibited solid returns even on recent data, i.e. in period when value strategy lagged behind the rest of the market. Eisfeldt & Papanikolaou (2013) also demonstrate that firms with high level of intangibles outperform its counterparts. On the other hand, evidence of Chan *et al.* (2001) does not support any link between R&D costs and future returns. In particular, three years after portfolio construction, stocks incurring R&D costs yielded higher average returns only by 0.15 p.p. than those not incurring any. Perversely, they found evidence that stocks with increased R&D are subject to elevated volatility.

Investment factor has already been covered above. However, this factor is also associated with a need for financing. Various sources of funds and financing structures affect cost of equity as it determines the volatility of cashflows attributable to shareholders. Bhandari (1988) shows that leverage, measured by debt-to-equity ratio, implies certain premium even when controlling for stock's β and size. This conclusion clearly indicates the failure of market β to fully proxy for such an obvious source of risk - leverage. This conclusion is supported by similar study of Dhaliwal *et al.* (2006). They show that direction of the relationship between leverage and cost of equity is consistent with theory and its magnitude depends on corporate tax level due to tax-deductibility of interest expenses. Deduction of interest expenses reduces the cost of debt to the company, further benefiting its shareholders.

While capital structure is a risk-based factor, the events of capital raising and their implications on future returns are recognized as rather behavioral anomalies. Examples of such fundraising events are IPOs, SEOs, bond issuance etc. Investing in IPOs is considered risky, but potential payout is high. Existing academic research indicates that risk of losing money overweights the potential benefits, especially in the long-run (see Ritter, 1991; Loughran & Ritter, 1995; Spiess & Affleck-Graves, 1995; Bradshaw *et al.*, 2006). Spiess & Affleck-Graves (1995) extends their study to SEO event to conclude that post-IPO underperformance is an anomaly associated with equity offering in general. After controlling for firm's age and B/M ratio as well as for specifics of the offering itself, they argue that on average firms after SEO tend to exhibit abnormal negative returns, especially in the long-run (see also Loughran & Ritter, 1995; Bradshaw *et al.*, 2006). Most of the mentioned studies offer explanation of managers exploiting insider information to sell equity when the stock is positively mispriced or artificially inflating earnings in the period preceding the SEO (Cohen & Zarowin, 2010). Totally reversed situation to selling firm's equity due to overvaluation is buying own firm's equity for undervaluation. It has become common practice that a firm buys back its own stock rather than paying cash dividends, e.g. in the last decade S&P500 companies consistently spent more cash on share repurchases than dividends. The usually pronounced reason by managers that it is a good investment is generally supported by existing research, i.e. it documents abnormal returns for share repurchasing firms in the long-run, especially in case of value stocks (Ikenberry *et al.*, 1995; Zhang, 2005).

Asset pricing (and other economic) models are usually a simplification of more complex reality, and so rely upon many unrealistic assumptions, e.g. full rationality of agents. Behavioural finance relaxes this naive assumption and applies psychology to understand financial market behaviour. Reversal anomaly has already been mentioned above as it serves as alternative explanation to value factor (Lakonishok *et al.*, 1994). The research pool on the idea of buying past losers and selling past winners includes Bondt & Thaler (1985) documenting that past losers over the last 3 years are set to become future winners over the next 3 years, Chopra *et al.* (1992) also seeing long-term mean reversal for prior losers in a 5-year span etc. However, Jegadeesh (1990) identifies only short-term reversal effect (1-month) and conversely finds positive autocorrelation for longer lags indicating rather momentum effect. Even though momentum and reversal strategies appear to be conflicting with each other, they do not always compete as the time frames of these effects do not necessarily overlap. Momentum strategies are profitable primarily in the period of 3-12 months following portfolio formation (Jegadeesh & Titman, 1993), which is consistent with most of the studies on mean reversals as these show reversal effect materializing either within shorter period (1 month) or longer period (3-5 years). Moreover, combination of these two seemingly rival strategies can yield higher abnormal returns than choosing one or the other (Kot & Chan, 2006). Existing literature identifies several proxies exhibiting predictive power on top of the standard returns-related proxy variables for reversal and momentum anomalies, such as 52-week high (George & Hwang, 2004), past performance of overall industry while controlling for individual firm characteristics (Moskowitz & Grinblatt, 1999) etc.

There are dozens of other anomalies detected by prior literature and related to frictionality of financial markets or psychological element of investing. Haugen & Baker (1996) emphasized the importance of stock liquidity as low liquidity comes with additional transaction costs, e.g. wide bid-ask spread (see Amihud & Mendelson, 1986). Amihud (2002) also suggests that stock returns reflect an illiquidity premium. Haugen & Baker (1996) measure liquidity with various volume-based metrics such as average daily volume-to-market capitalization ratio, 5-year trend of this ratio etc. Further liquidity measures include high-low estimator [5] or turnover ratio [6] (Leirvik *et al.*, 2017). Past trading volume is also linked to momentum and reversal strategies. Specifically, stocks with high turnover ratios exhibit faster reversal of momentum effect, translating into lower future returns than low-volume stocks (Lee & Swaminathan, 2000). Seasons of the year also apparently affect investor's behaviour. Particularly, substantially skewed returns could be observed in January, especially for small capitalization stocks and past losers (Thaler, 1987; Keim, 1983). Past losers being usually a subject to January effect is consistent with the hypothesis of tax-loss harvesting at fiscal year-end (Poterba & Weisbenner, 2001).

[5] High-low estimator = (daily price high - daily price low) / daily price high

[6] Turnover ratio = no. of shares traded at a given day / no. of shares outstanding

2.4 Predictability of Stock Returns

Existing pool of literature on asset pricing is flooded with factors and anomalies explaining variation in expected returns. Between 1970 and 2010 prior research identified at least 330 firm-specific variables associated with stock returns (Green *et al.*, 2013). How many of those variables really carry an explanatory or predictive power is questionable due to various types of biases detected across the academic papers.

Several authors point out the existence of *survivorship bias* in previous research (see Kothari *et al.*, 1995; Brown *et al.*, 1995, etc.). Survivorship bias is a specific type of more general bias - selection bias. It arises when inactive or delisted companies are omitted in the sample, e.g. companies delisted due to bankruptcy, merger or going private. As a result, financial performance of the sample might be positively biased. Especially in case of analysis of certain anomalies such as B/M ratio or leverage, the investigated effects might be substantially distorted. Companies with these characteristics usually face increased likelihood of bankruptcy or delisting, e.g. Kothari *et al.* (1995) found that the value effect proxied by B/M in Fama & French (1992) is less economically and statistically significant than they reported.

Look-ahead bias occurs when information, that had not been known in a given period, is taken into account in the analysis, e.g. forming portfolios in January based on last calendar year accounting data probably produces look-ahead bias. In such case the accounting figures probably were not announced at the time of the portfolio decision. Annaert *et al.* (2002) document on their sample of European stocks that correcting for look-ahead bias yields only insignificant value premium of 2%, as opposed to significant premium of 11% for sample suffering for the bias. Another potential distortion of analysis outcome occurs in the process of evaluating portfolio performance against an index. Compositions of benchmark indices such as S&P 500 is subject to performance and compliance with certain characteristics for member companies. Thus the composition is somehow dynamic through time. Daniel *et al.* (2008) report that testing portfolio performance against S&P 500 might produce bias up to 8% per annum, when end-of-period index composition is used instead of the composition corresponding to the time of portfolio formation.

In many cases, factors (and associated hypotheses) compete with each other, e.g. value considered as one of the most robust factors explaining returns (see Fama & French, 1993; 2015; Chan & Chen, 1991, etc.) is challenged by overreaction anomaly (see Lakonishok *et al.*, 1994; Bondt & Thaler, 1985, etc.). Stocks with high B/M ratios are most likely past losers, hence overreaction explanation is also relevant in this case. Correlation between these two variables will most likely be present. Chan *et al.* (1991) documents significant drop in explanatory power of B/M ratio when regressed along with past loser-winner proxy. Similar logic applies also to size factor and liquidity-related anomalies. In five-factor model (see Fama & French, 2015), accounting for profitability and investment absorbs the explanatory power of value effect. Analogous findings that a factor subsumes the effects of the others are common in asset pricing literature. Such subsumation is the main reason for simplicity of benchmark asset pricing models, i.e. CAPM, three and five-factor models. Including only the most robust factors in regression models reduces the risk of multicollinearity and potential overfitting. Especially for task of prediction of expected returns, the risk of overfitting is for traditional methods one of the key limitations to account for all the discovered factors and anomalies, e.g. inclusion of hundred of predictors led to zero predictability of future returns with negative values of R^2 in Gu *et al.* (2020). Lewellen (2015) shows that out-of-sample performance is similar for 3 and 15-predictor specifications. As a result, only handful out of the hundreds of potential candidate variables (see Green *et al.*, 2013) is usually selected. Further pitfalls stem from the need to specify the functional forms upfront, including interactions between predictors. Even though prior literature provide guidance in this regard, the set of potential specifications is large and the correct form is always ambiguous.

Recent research aimed at prediction of expected returns more frequently adopts machine learning algorithms such as random forests, gradient boosting machines or neural networks to cope with the limitations of Ordinary Least Squares (OLS) regression. Nowadays, these tools are widely utilized across finance industry, especially for purposes of portfolio selection, risk management and short-term trading. Machine learning techniques are well suited for the task of prediction of expected returns as they mitigate the risk of overfitting and remove the need to identify complex functional forms and interactions terms manually. Moreover, arbitrary set parameters for controlling the fitting process allow for inclusion of substantially greater number of predictors. In the environment of hundreds of reported anomalies (e.g. Hou *et al.* (2020) counts over 450 identified by prior research), selection of the right predictors at the discretion of the researcher is not effective. Mclean & Pontiff (2016) demonstrate that the majority predictive power of reported anomalies disappears post publication anyway. Generally, key drawback of machine-learning-based methods is in individual effects interpretation. However, for the sole purpose of the prediction task, these feature is not necessarily required. Estimation of individual effects of the predictors is still possible but its accuracy is subject to model complexity (see Friedman, 2001). Gu *et al.* (2020) emphasize the superiority of machine-learning algorithms relative to OLS-based models in the field of equity risk premium prediction. They report substantial improvement in out-of-sample R-squared when random forests or neural networks are applied, especially for large capitalization stocks. Their long-short value-weighted portfolio based on neural networks delivers Sharpe ratio of 1.35, while three-factor (size, M/B and momentum) OLS-based strategy has Sharpe ratio of 0.61.

3 Data & Universe Construction

3.1 Complete Sample

We acquired market, accounting and I/B/E/S data from Refinitiv Eikon Datastream for all non-financial[1] firms listed on stock exchanges 1) that are located in North America or Europe and 2) with market capitalizion exceeding USD 1 trillion, i.e. NYSE, NASDAQ, NASDAQ Nordic, Euronext, Toronto SE, SIX Swiss Exchange and Deutsche Boerse. For European stocks all fundamental and market data are automatically converted to USD using historical exchange rates provided by Refinitiv. All the categories of data are retrieved on monthly frequency for the period from March 2000 to October 2021, spanning more than 21 years. Please see cross-sectional composition of stocks in the whole sample in Table 3.1. You can observe that the sample is geographically well balanced. Though, the split is not equal.

Table 3.1: Cross-section of stocks in the complete sample

	min	mean	max
Europe	1,731	2,740	3,423
North America	2,297	3,204	4,370

Monthly number of stocks in the entire sample between March 2000 and October 2021. The number fluctuates over the observed period due to new listings and delistings.

The dataset also includes stocks that were delisted during the observed period in order to eliminate *survivorship bias*. There is total of 2,114 such securities in the sample. We treat the delisted stocks by excluding them from the data one month prior its delisting date. This type of treatment is primarily due to

[1] This thesis defines industries and sectors according to North American Industry Classification System (NAICS). The companies that fall into category *Finance and Insurance*, as defined by NAICS, are considered financial and excluded from the data.

lack of available information regarding reasons[2] of the individual delistings on Reuters Eikon Datastream.

The frequency and time alignment of the accounting data is designed such that it reflects the real information set available at the time of each portfolio rebalancing data. For last day of each month, we retrieve last twelve months (LTM) data aligned according to financial results announcement dates (as opposed to usually used fiscal period-end dates). E.g., On 5 November, 2020 *Discovery, Inc.* submitted its 10-Q form for fiscal quarter ended 30 September, 2020. Thus, the retrieved accounting data for *Discovery*, Inc. for September and October 2020 does not contain the figures from the most recent fiscal quarter as it has not been published yet. Finally, in observation for November 2020 the LTM accounting figures reflect also the quarter ended 30 September, 2020. Such data structure allows for mimicking real-world situation an investor faces at the time of his decision-making process. This not only mitigates the risk of *look-ahead bias* but also ensures that the portfolio construction process is based on the most recent public information available each month.

Individual accounting figures that appear unreliable due to a conflict with accounting standards are set to missing. E.g., positive cost items or negative cash dividends paid. On top of that, firm-year observations with apparently illogical or highly anomalous values in essential accounting variables are dropped from the sample. This covers long-term debt greater than assets, profit margins greater than 100%, or deeply negative etc.

In addition to the stock-level data, we obtained monthly size, value, profitability, and investments factor premiums along with market and risk-free returns for developed countries from Kenneth R. French's data library[3]. These factors are based on global developed markets as defined by the source. As a proxy for credit cycles serves ICE BofA US High Yield Index Option-Adjusted Spread[4] from FRED database. It is calculated as differences between US option-adjusted high-yield index (capitalization-weighted index composed of public

[2]Common reasons for removal from a stock exchange are failure to meet listing requirements, voluntary delisting, or getting acquired by another company.

[3]`http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

[4]Ice Data Indices, LLC, ICE BofA US High Yield Index Option-Adjusted Spread [BAMLH0A0HYM2], retrieved from FRED, Federal Reserve Bank of St. Louis; `https://fred.stlouisfed.org/series/BAMLH0A0HYM2`

USD-denominated below investment grade bonds) and spot treasury curve. Another proxy for uncertainty prevailing in financial markets are average monthly values of VIX Volatility index, which represents expected volatility for S&P 500 in the next 30 days as implied by current call and put options.

3.2 Universe Construction

In this section, the entire sample is trimmed into an universe composed of only stocks with lower absolute valuations, cheap relative valuations and high level of financial leverage. This universe is supposed to include stocks with mispricing potential and with exposure to premium-yielding risk factors. This thesis builds upon the findings of Rasmussen & Chingono (2015), so our definition of the universe is equivalent to theirs. We construct such set of companies by applying three separate filtering conditions to all firms in the sample. The conditions are defined as follows:

1) **Stocks with market value of equity between 25^{th} and 85^{th} percentile, sorted each month.** Definition of *small-capitalization* stocks is usually arbitrary in academic literature and sometimes differs from the ranges seen in practice, e.g., Fama & French (2008) use USD 2.3 billion as the upper threshold for small stocks, while Nasdaq stock exchange considers small-capitalization stocks as those with market value of equity less than USD 1 billion. Since this thesis defines the thresholds using percentiles each individual month, the thresholds are variable in time. This selection criteria is also aimed to exclude micro-capitalization stocks which might lack sufficient liquidity.

   In Table 3.3 you can see structure of the universe in terms market capitalization and other selected variables for pooled data (cross-sections across all periods pooled together) and for cross-section in September 2021, the last month of the observed period. Note that it shows only such companies that meet all three criteria.

2) **Stocks with *book-to-market* ratio above its median value - cheaper half of the sample every month**. This criteria intends to further restrict the universe to stocks with characteristics that historically demonstrated a risk premium. Value premium usually compensates for financial

distress (Fama & French, 1993) which also creates further space for uncertainty and mispricing by investors. Detection of mispriced stocks among high-value stocks would represent additional source of return on top of historically exhibited value premium.

In prior literature, various proxies for value are used, e.g., *B/M ratio, E/P ratio, EBITDA/EV* (see Fama & French, 1992; Davis, 1994; Rasmussen & Chingono, 2015, and others) etc. This study adopts *B/M* metric as it is the most conventionally used in academic literature and it is known for most significant relationship with returns. Descriptive statistics are presented in Table 3.3.

3) **Stocks with leverage (defined as Long-term debt/EV) higher than 50^{th} percentile of the sample, each month.** Even though B/M effect is supposed to already contain risk associated with financial leverage, this condition ensures its complete presence in the resulting universe.

Above described breakpoints for the individual criteria for the universe construction are set arbitrary such that sufficient number of observations is passed through the filter each period. Minimum, maximum and average number stocks in cross-section are listed in Table 3.2. Naturally, observations with missing values in variables essential for the universe selection are automatically dropped. E.g., when it is not possible to calculate *Long-term / EV* due to a missing component, the observation will not pass through the universe selection criteria. In addition, observations with outliers[5] or missing values in *1-month-ahead returns* variable are completely removed from the universe.

Table 3.2: Cross-section of stocks in the universe

	min	mean	max
Europe	36	325	464
North America	127	379	570

Monthly number of stocks in the universe of leveraged value small-caps between March 2000 and September 2021.

As a result, there is total of 182,526 firm-month observations comprising 3,793

[5] In this thesis Tukey's method (see Tukey, 1977) for outlier detection is adopted. According to this rule, observations lying outside range $[Q1 - 1.5(Q3 - Q1), Q3 + 1.5(Q3 - Q1)]$ are considered outliers.

unique stocks over 259 months starting on March 2000 and ending on September 2021. The amount of observations per month in the universe rises in time substantially, especially in early months of the observation period. This happens primarily due to increasing availability of data for essential accounting variables on Refinitiv Eikon Datastream, especially for stocks listed in Europe. E.g., in March 2000 only 10% of the European listed stocks in the sample had non-missing value for *Total Assets* and other fundamental figures. However, this improves rapidly in the early 5 years of the observation period. Since 2005 it is consistently over 60% of European stocks in the complete sample with sufficient non-missing figures in essential variables. In Table 3.2 you can see that the universe is on average geographically balanced.

Selected individual characteristics for the firm-year observations, that are allowed to the universe, are presented in Table 3.3. Panel A summarizes the data pooled together disregarding any effects of time. On the other hand, Panel B is attached primarily for convenience as some characteristics developed through the observed 21 years (e.g., average market capitalization nearly doubled during the 21-year period). Since the pooled statistics are more relevant for the methodology of this paper, concerning Table 3.3 it is always referred to statistics in Panel A, unless stated otherwise. It is important to note that most of the variables presented in the table are ratios susceptible to extreme values in case of close-to-zero values in denominator.

Given the skewness and kurtosis, out of the presented variables *1-month returns (%)* has probability density function that is closest to normal. With positive mean at 0.8% and standard deviation of 8.6% its distribution is defined as symmetric and slightly leptokurtic. It is consistent with previous research regarding distribution of stock returns (see Hwang & Satchell, 1999; Kim & White, 2004). Though it should be kept in mind that outliers in this variable are stripped off. In terms of market capitalization, we can see that majority of firms exceed USD 500 million valuation mark. Standard deviation as much as USD 1.3 billion is attributable to several micro-capitalization stocks still appearing in the Universe. As implied by Table 3.3, average stock in the universe is valued at 0.88 multiple of book value and 16.7 multiple of earnings. Average enterprise value is 6.3 times EBITDA. We can also observe that EBITDA/EV valuation is the most relatively dispersed valuation metric in the table and also the one with fattest tails. Comparing the valuations from Panel A to Panel B, stocks became apparently more expensive during the observed period, particularly with respect to book value and earnings. Solvency, another filtering criteria for the universe, stands on average at 38% long-term debt relative to EV with standard deviation of 23 percentage points. Average firm-year observation has operating profit 4.2 times the annual cash outflows for debt servicing, however, median value of 1.4 is more representative value for average constituent of the universe. Asset growth was positive in the last twelve months to September 2021 in more than 50% cases (see Panel B) although this set of firms targets the distorted ones. Mean value of 98% growth in assets (see Panel A) is affected by extreme outliers as suggested by kurtosis of more than 46,000. Profitability at the level of EBITDA is about 20%, whereas fatter left tail of its distribution suggests that loss-making companies are present in the sample.

Table 3.3: Descriptive statistics for the universe

Panel A: Pooled observations

	Q1	Median	Mean	Q3	Standard Deviation	Skewness	Kurtosis
1-month return (%)	-4.39	0.66	0.83	6.03	8.59	0.04	3.1
Market cap (USD mil.)	201	545	1,079	1,432	1,323	2.2	8.6
Book-to-Market	0.60	0.81	1.14	1.19	2.31	21.4	581
LT debt/EV	0.22	0.33	0.38	0.49	0.23	3.2	35.4
EBITDA/EV	0.09	0.13	0.16	0.19	0.58	-172.1	59,898
E/P	0.03	0.07	0.06	0.11	0.40	-1.0	987
ROA (CFO)	0.04	0.07	0.07	0.10	0.06	-0.1	12.1
DSCR	0.43	1.43	4.17	4.52	875.08	315.5	120,647
Asset growth	-0.02	0.05	0.98	0.15	136.35	204.3	46,385
EBITDA margin	0.08	0.13	0.20	0.24	0.20	1.4	5.9

Panel B: Cross-section in September 2021

	Q1	Median	Mean	Q3	Standard Deviation	Skewness	Kurtosis
Market cap (USD mil.)	411	1,089	1,915	2,747	2,078	1.6	4.9
Book-to-Market	0.48	0.69	0.87	0.97	1.77	27.8	845.1
LT debt/EV	0.23	0.34	0.39	0.49	0.21	1.1	4.6
EBITDA/EV	0.07	0.12	0.13	0.17	0.13	2.4	19.6
E/P	0.01	0.06	0.04	0.10	0.23	7.0	175.1
ROA (CFO)	0.04	0.07	0.07	0.11	0.07	0.3	10.2
DSCR	0.24	1.15	2.89	3.53	30.46	0.00	168.7
Asset growth	-0.01	0.06	0.10	0.12	0.39	12.7	235.3
EBITDA margin	0.09	0.14	0.21	0.28	0.22	0.9	5.1

Panel A presents descriptive statistics for selected variables for all firm-year observations in the universe of leveraged value small-caps, i.e., there is no time or cross-sectional discrimination. The universe totals 182,526 firm-year observations comprised of 3,793 unique stocks and their characteristics over the period from March 2000 to September 2021. Note that observations outlying in variable *1-month return (1%)* are dropped. Panel B shows the statistics only for cross-section as of September 2021. Precise definitions of the variables are disclosed in Appendix A.

3.3 Variables

As already described above in Section 3.1 and demonstrated on *Discovery, Inc.* example, the accounting data is retrieved in LTM format and aligned according to announcement dates of the most recent quarterly (semi-annual, in case of European firms) results. Naturally, LTM format affects only flow variables. Stock variables such as balance sheet data are simply the most recently published.

With reference to extensive pool of previous studies focused on factors and anomalies affecting stock returns in the cross-section, we define 42 stock-level features. These include various valuation metrics; profitability, solvency and liquidity ratios; trading characteristics and forward-looking estimates by analysts. Variables expressing a relative change or a difference between accounting figures in time are defined on year-over-year basis, unless specified otherwise. Detailed description of individual variables adopted by this study is presented in Appendix A.

4 Methodology

In this section, we discuss the methods applied in individual steps of our analysis, which encompasses 1) forecasting the ability to pay down long-term debt; 2) estimation of future returns and subsequent one-way portfolio sorts construction; 3) testing of mispricing strategy based on the portfolio sorts; and 4) examination of the mispricing strategy with respect to credit cycles and market volatility. Unless otherwise stated, the methods presented in this section are applied only to the universe of leveraged value small-caps (as defined in Section 3.2), not the entire sample. Moreover, the dataset is further split into three separate time frames - training sample, validation sample and testing sample. *Traning* is the longest one (March 2000 - March 2013) and represents the main component of in-sample period. Validation sample immediately follows (April 2013 - March 2015) and serves for purpose of hyperparameter tuning. Once the arbitrary hyperparameters are determined, final model is fitted using union of training and validation samples (=the whole in-sample period) and tested on *testing* sample (=out-of-sample period) spanning from April 2015 to September 2021.

4.1 Forecasting Future Debt Reduction

Since Rasmussen & Chingono (2015) argue that deleveraging is the main driver of future returns for leveraged value small-capitalization stocks, first step in our analysis is to estimate future reduction in long-term debt. Considering that debt is reduced mainly by means of sufficient cash generation from business activities, we use fundamental-only stock-level lagged characteristics as the predictors. Behavioral signals are not taken into account. Specifically, we estimate the ability to pay down long-term debt in the period of next twelve months (although data periodicity is monthly). Since companies in the sample usually report their interim results quarterly or semi-annually, 1-month-ahead estimation of the debt reduction is not applicable. To summarize this, we forecast the ability to reduce debt exposition in the next twelve months using LTM accounting-based metrics, EBITDA/EV multiple (to proxy for negative fundamental information that cannot be captured by accounting figures) and forward-looking analysts estimates. All the predictive signals entering the estimation reflect the most recent public information. Generally, we estimate the following equation:

$$y_{i,t+12} = f(x_{i,t,1}, x_{i,t,2}, .., x_{i,t,K}) + \epsilon_{i,t} \tag{4.1}$$

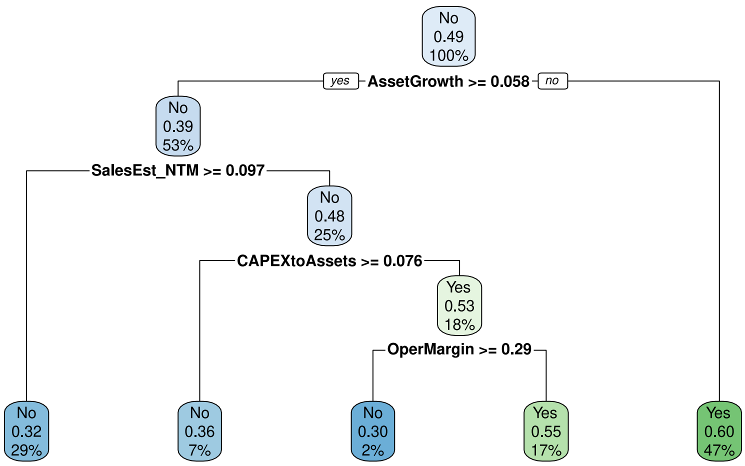
where $y_{i,t+12}$ is binary variable taking value of 1 if the company i reduces its long-term debt in next twelve months (i.e. long-term debt as of *t+12* is less than in period t) and 0 otherwise. t denotes individual monthly periods. x_k denotes individual predictors. Full list of such variables and their definitions are listed in Appendix A.

For such classification problems (that is using firm-level characteristics to model future binary events, e.g. bankruptcies), wide range of methods is available and thoroughly tested in previous studies. Taking extensive literature on predicting the binary event of bankruptcy as model performance benchmark (see Gepp *et al.*, 2010; Barboza *et al.*, 2017), this thesis utilizes superior predictive power of machine-learning classification methods over the conventional ones such as logistic regression. The advantages of machine-learning techniques include accounting for potential variable interactions not being limited by loss of degrees of freedom, allowance for non-linear relationships, and arbitrary over-fitting control mechanisms. The mostly referred to drawback of limited intepretability does not pose a problem in case of our application and we provide interpretation of individual effects anyway (see Section 4.3 for greater detail).

4.1.1 Gradient Boosting Machines with Trees (GBM)

Decision trees are one of the most favourite machine-learning algorithms due to to their simplicity in terms of both application and interpretation. Most straightforward way how to briefly introduce the decision tree algorithm is by a graphical illustration. See depicted one in Figure 4.1. Each split in the decision tree is made such that a selected loss function is minimized. At the same time, each splitting is subject to arbitrary conditions that serve to mitigate overfitting (e.g., minimum number of observations per split or maximum number of splits). Classification trees usually minimize Gini impurity measure (which is the probability of misclassifiying an observation) to determine where and whether to make a split.

Figure 4.1: Example of a decision tree



Decision tree with target (dummy) variable whether the company will be able to reduce its debt in the next twelve months. The first split (i.e. the upper one) is based on predictor *AssetGrowth* and implies that for values lower than 0.058 the observation is classified as "Yes", i.e. the firm will be able to reduce its debt, otherwise it is subject to further splitting down the tree. This decision tree has 5 terminal nodes and depth of 4 (i.e. there is maximum of 4 consecutive splits).

"Boosting" is an ensemble method that combines several underlying models into single and more precise one. Particularly, gradient boosting decision trees is a combination of multiple decision trees, "weak learners", via gradient descent optimization technique into a "strong learner". However, the individual trees - that are combined - are not independent of each other. It is a sequential process where individual weak learner is fitted with respect to the errors (defined by a selected loss function, which is Bernoulli loss in our case) of the previous one. Generally, stepwise implementation of a GBM algorithm of Friedman (2001) with M iterations, loss function $L(y, \widehat{F})$, input data (x_i, y_i) with $i = 1, ..., N$ and weak learner $h(x, \theta_m)$ is as follows:

1. Initiate $\widehat{F}_0$ with a constant such that loss function $L(y, \rho)$ is minimized, and

2. for each iteration m, $(m = 1, ..., M)$:

   (a) Calculate negative pseudo residuals, which are defined as

   $$\widetilde{y}_i = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)} \text{ for } i = 1, ..., N.$$

   Note that the pseudo residuals would equal to actual residuals if the selected loss function was squared-error loss L_2.

   (b) Fit the weak learner $h(x, \theta_m)$ on the pseudo residuals and input data $(\widetilde{y}_i, x_i)$ for i = 1,...,N, i.e.

   $$\theta_m = arg \min_{\theta,\beta} \sum_{i=1}^{N} [\widetilde{y}_i - \beta h(x_i, \theta)]^2$$

   (c) Find the optimal gradient descent step-size (i.e. optimal weight of the m^{th} weak learner in the final prediction function) by minimization of the specified loss function.

   $$\rho_m = arg \min_{\rho} \sum_{i=1}^{N} L(y_i, \widehat{F}_{m-1}(x_i) + \rho h(x_i, \theta_m))$$

   (d) Update the prediction function as

   $$\widehat{F}_m(x) \leftarrow \widehat{F}_{m-1}(x) + \rho_m h(x, \theta_m)$$

.

For further detail on the algorithm and further discussion the use of various loss functions and weak learners, please refer to seminal paper Friedman (2001).

Hyperparameters are model characteristics that determine how the model learns. It needs to be specified before the model sees its "learning materials" - the training data. *Shrinkage* hyperparameter known as learning rate determines contribution of each tree to the resulting function. *Number of trees* determines the number of trees to be grown - iterations used for the optimization of the selected loss function. Third important hyperparameter of the algorithm is *Maximum tree depth* specifying maximum number of splits per individual tree. High values of *Number of trees* and *Shrinkage* typically provide better fit in-sample, however at the expense of out-sample stability. More trees increases the risk of overfitting. On the other hand, lower *Shrinkage* diminishes the importance of the errors that are fitted by every next tree, which decreases the risk of overfitting.

Chingono & Rasmussen (2016) served as a benchmark in the matter of hyperparameter setting, however, we still performed extensive grid search considering values {400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400} for *Number of trees*; {0.1, 0.05, 0.01, 0.005} for *Shrinkage*; {5,7,10} for *Maximum tree depth* and {5,10} for *Minimum observations per node*. In this tuning process, all the individual model specifications were trained on training sample and subsequently tested on validation sample. Set of hyperparameters with maximum accuracy on the validation sample is the preferred here; i.e. 2000 trees learning at rate of 0.05 with maximum of 7 consecutive splits per tree and minimum of 5 observations per terminal node. Once the hyperparameters are optimized, final model is fitted on union of the training and validation samples.

In case of classification, the final arbitrary parameter required to be specified in before out-of-sample model evaluation is probability threshold. We set the probability threshold based on Youden's J statistic (see Youden, 1950) using validation sample, i.e., we set the threshold such that sum of *specificity* and *sensitivity* is maximized.

4.2 Mispricing Strategy

We start with forecasting future *1-month-ahead returns*. Mispricing strategy is then executed by forming long-short portfolios each month based on the forecasted future performance of individual stocks.

4.2.1 Predicting Future Returns

Analogously to Section 4.1, gradient boosting ensemble of tree-based learners is the method of our choice. Specifically, it is gradient boosting regression trees algorithm with Gaussian L_2 loss function, which has been thoroughly tested by prior studies in the exercise of predicting cross-sectional returns. It has also been demonstrated to be advantageous over conventional regression methods, given adequate sample size (see Gu *et al.*, 2020; Leung *et al.*, 2021; Tobek & Hronec, 2021; Choi *et al.*, 2021).

34 stock-level characteristics are supplied to the supervised learning process in

order to predict *1-month-ahead returns*. Additionally to accounting ratios and I/B/E/S estimates used for the prediction of future debt reduction, we extend the set of predictors for behavioral factors such as momentum and short-term reversal, various valuation multiples and the already predicted future ability to pay down debt. To summarize, the following equation is estimated

$$r_{i,t+1} = g(\hat{y}_{i,t+12}, x_{i,t,1}, x_{i,t,2}, .., x_{i,t,S}) + \epsilon_{i,t} \tag{4.2}$$

where $r_{i,t+1}$ represents total return of stock i for period *t+1*, $\hat{y}_{i,t+12}$ denotes the already predicted probability that stock i reduces its long-term debt in next twelve months. x_s denotes individual lagged predictors. Full list predictors is disclosed in Appendix A.

Study Gu *et al.* (2020) serves as benchmark in terms of hyperparameter tuning, so we center our grid search around their referenced values. The hyperparameters are optimized throughout all combinations of the following values {200, 400, 600, 800, 1000, 1200, 1400, 1600, 1800, 2000, 2200, 2400, 2600} for *Number of trees*, {0.1, 0.05, 0.01, 0.005} for *Shrinkage*; {5,7,10} for *Maximum tree depth* and {5,7,10} for *Minimum observations per node.* Model is trained on training sample and parameters subsequently selected such that root mean squared error (RMSE) on Validation sample is minimized. Error-optimizing hyperparameter combination is growing 200 trees at leaning at rate of 0.01 with maximum tree depth of 10 and at least 5 observations per leaf. Final tuned model is trained on union of training and validation samples.

4.2.2 Portfolio Construction and Testing

Each month we sort the universe based on the *1-month-ahead returns* variable predicted in Subsection 4.2.1. The sorted stocks are then split into deciles, e.g. top decile portfolio consists of 10% highest-ranked stocks in terms of the predicted returns. Each month the number of firms in the portfolios differs as the size of the universe is also period-dependent. Finally, at the beginning of each month we construct long-short portfolio by buying stocks in the top decile and selling short stocks in the bottom decile. The constructed long-short portfolio is therefore rebalanced based on the most recent information every month. The strategy and generated returns assume 1/3 of the initial dollar exposition is held in cash, e.g. we go long stocks worth \$1, short stocks worth \$1 and hold

\$1 in cash. Transaction costs, cost of selling short and returns on cash holdings are not taken into account. As a result, each month we are supposed to have a portfolio with low exposure to market, size and value factors that seeks to exploit idiosyncratic returns.

Performance of the long-short portfolio is tested against benchmark portfolio and risk factors commonly adopted for testing purposes in asset pricing literature, i.e., CAPM and multifactor models of Fama & French (see Equation 2.1, Equation 2.5 and Equation 2.6). Testing period comprises 77 months starting April 2015, ending September 2021.

The hypothesized mispricing strategy assumes that weak, or semi-strong market efficiency do not hold, and seeks to exploit failures of market to correctly reflect available information into prices. Consequent hypothesis conjectures that in periods of increased uncertainty, investors might be more susceptible to misprice the securities, especially those that are exposed to various risk factors such as value, size and leverage. To test this hypothesis, we regress *Jensen's alpha*[1] to proxy variable for credit cycles. Analogously, we regress the abnormal returns (= Jensen's alpha) to market volatility represented by VIX Volatility index prices. The following equations are estimated

$$\hat{\alpha}_t = \gamma + \beta_1 OAS_t + \epsilon_{i,t} \tag{4.3}$$

$$\hat{\alpha}_t = \gamma + \beta_1 VIX_t + \epsilon_{i,t} \tag{4.4}$$

where $\hat{\alpha}_t$ denotes *Jensen's alpha* in period t. Its calculation is presented in Equation 4.5 below. OAS_t denotes average option-adjusted spread between below investment-grade US dollar-denominated bonds and spot treasury curve during month t.

$$\begin{aligned} \hat{\alpha}_t &= r_t - E(r_t) \\ \hat{\alpha}_t &= r_t - [r_{rf,t} + \hat{\beta}(r_{M,t} - r_{rf,t})] \end{aligned} \tag{4.5}$$

where r_t is realized total return of the selected long-short portfolio in month t, and $\hat{\beta}$ is regression coefficient already estimated using CAPM as desribed in second paragraph of this subsection. Analogously, to this CAPM-based

[1] *Jensen's alpha* is a risk-adjusted measure of portfolio performance derived from CAPM in Jensen (1968) and applied to measure performance of mutual funds.

calculation, we also calculate *Jensen's alpha* using thee and five-factor models by Fama & French and estimate the corresponding regressions (Equation 4.3).

4.3 Interpretability Measures for GBM

Machine-learning techniques including the supervised ones are sometimes touted as "black box" methods. This is primarily due to its limited interpretability. While assessment of the accuracy of the model predictions is equivalent to conventional methods, interpretation of individual effects and their contribution to those estimates requires more complex tools.

One of such tools usually reported is relative importance of individuals signals. You can find charts depicting importance of individuals signals in Appendix A. The relative importance is often reported in scaled values. However, we disclose values not standardized so that further elaboration on top of solely relative influence is possible.

For the purpose of the classification problem, we measure the importance of individual predictors using *permutation test* of Breiman (2001). This methodology determines the influence of a predictor x_k as the decrease in classification accuracy when variable x_k is "noised up". Particularly, the decrease in classification accuracy is calculated as the difference between accuracy of the fitted GBM model on original training data and the classification accuracy of the GBM model such that for each tree the values of variable x_k are randomly permuted (while values of other predictors remain unchanged). The predictors with largest increase of the missclasification rate are deemed most important.

In this study, for regression-tree-based models, relative importance is denoted by the improvement in squared errors attributable to individual signals as described by Friedman (2001). On the level of individual trees the original study approximates absolute improvement associated with an input variable as follows:

$$\hat{I}_j^{\,2} = \sum_{t=1}^{J-1} \hat{i}_t^{\,2} 1(v_t = j) \tag{4.6}$$

where j denotes the input signal of our interest, t denotes the individual nodes of the tree with J representing the terminal one. i_t^2 is the reduction in squared error, but taken into account only if the variable j is the splitting one. Subsequently, we get estimated reduction in squared errors for the entire model by taking average over all the trees in the GBM. E.g., for input variable *6-month cumulative return* on the right-hand side of Figure B.1 the original unscaled relative influence measure, $\overline{I_j^2}$, was 681,237; representing the overall reduction in squared error attributable to any splitting based on the given variable per average tree. We can further elaborate and divide the total reduced squared error by the total weight of the data (number of observations in our case) to estimate that inclusion of the particular variable decreases MSE of the model by 6. Such an absolute importance value should be treated with a caution due to interactions of the predictors, e.g., dropping a predictor that is usually the splitting variable in the upper part of the tree would affect the estimated MSE reduction attributed to the predictors closer to the terminal nodes. These variable importance metrics are also helpful for signal selection.

Though relative variable influences still do not provide any explanation how the input signal affects the target variable in terms of direction and magnitude. Fortunately, *partial dependence plots* is effective visualisation tool depicting the relationship of a signal and the target variable while averaging out the effects of the other predictors. Estimation of partial dependence for tree-based models was defined in Friedman (2001). For a set variables of interest z_u partial dependence function can be estimated by

$$\overline{F}_u(z_u) = \frac{1}{n} \sum_{i=1}^{n} \widehat{F}(z_u, z_{i,v}) \tag{4.7}$$

where $\widehat{F}$ denotes prediction function, i denotes observations in the training sample of size n and $z_{i,v}$ are the actual values of z_v in the training data. Stepwise practical implementation for a single variable of interest x with values $\{x_1, x_2..., x_l\}$ is as follows:

1. Duplicate the entire training sample l times, such that you create identical samples $\{S_1, S_2, ..., S_l\}$;
2. for $S_i (i = 1, 2, ..l)$
   (a) replace all values of variable x with value x_i and keep the other variables unchanged,

(b) calculate predicted value for each observation in the sample using prediction function $\widehat{F}$,

(c) calculate mean to get $\overline{F}(x_i)$;

3. plot the $\overline{F}(x)$ for $\forall x \in \{x_1, x_2, ...x_l\}$ (Greenwell, 2017).

Please see example of smoothed[2] *partial dependence plots* in Figure B.2

[2]We smooth the partial dependence plot using method of Local Polynomial Regression Fitting also know as "loess", where the fitting at each point x of the independent variable is determined using adjacent observations, weighted by their distance from x. For greater detail of this methodology see `https://rdrr.io/r/stats/loess.html`.

5 Results

5.1 Debt Paydown

In this section, we discuss the performance of the gradient boosting decision trees model (see detailed methodology in Section 4.1) trained on data between April 2000 and March 2015 and tested in subsequent 6,5 years - data the model has never seen before. Both in-sample and out-sample contain only observations from the universe, i.e., leveraged value small-caps. Table 5.1 displays confusion matrix whether long-term debt is reduced or not in the next twelve months, represented by two mutually exclusive classes "Yes" and "No". The probability threshold is set to 0.48 (which is the threshold that maximizes the sum of specificity and sensitivity on the validation sample, for further detail see Subsection 4.1.1), so observations with predicted probabilities greater than or equal to 0.48 are classified as "Yes" - those reducing its long-term debt in next twelve months - and "*No*" otherwise.

Table 5.1: Confusion matrix for Future debt paydown

		Actual outcome		
		No	**Yes**	**total**
Predicted outcome	**No**	19,170	8,824	27,994
	Yes	15,257	20,019	35,276
	total	34,427	28,843	

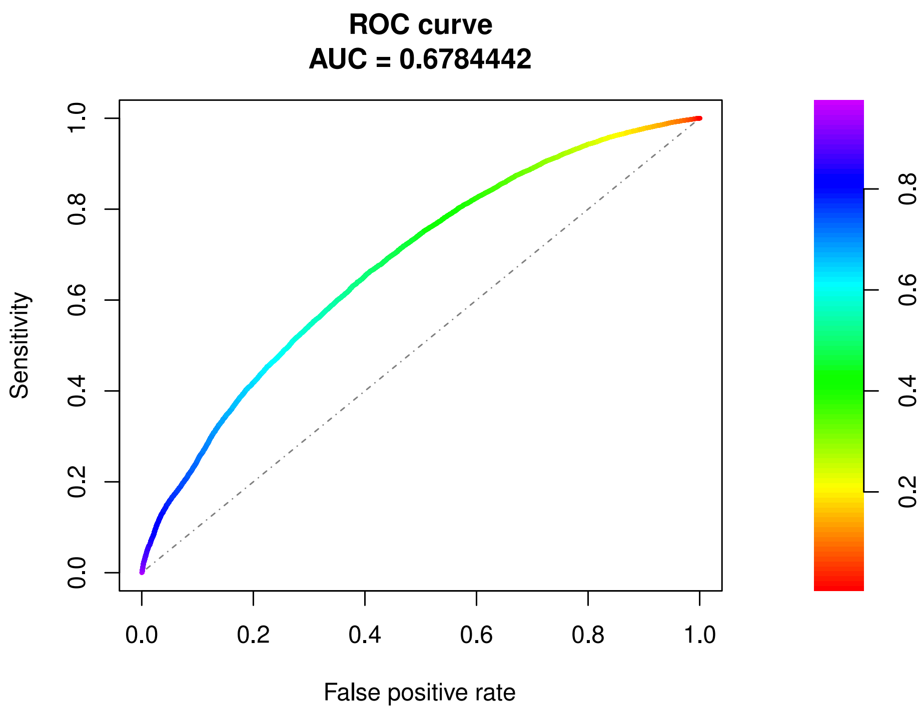
The matrix compares outcomes predicted by our gradient boosting machines model to the actual values of long-term debt reductions in the next twelve months. The predictions are for out-of-sample period (April 2015 - September 2020) for the universe of leveraged value small-capitalization firms.

First, we can observe that the sample is well-balanced since the prevalence of reducing debt is approximately 46%, not far from an equal split. However, the GBM algorithm predicts a slight majority of the firms to pay down the debt. So the algorithm tends to slightly favor positive outcomes. *Sensitivity* of 69% shows how successful the model is in detecting firms that actually reduce their debt. There is usually a trade-off between *sensitivity* of a classifier and its *precision*, which denotes how often the classifier is correct when it predicts a positive outcome. The best scenario is to get these two measures as high as possible, which is 100% for both measures. *F1 score* combines both measures such that we can evaluate the model performance regarding positive outcomes while taking the trade-off into account. F1 score for the GBM model amounts to 0.62. Rasmussen & Chingono (2015) argue that the ability to deleverage is the main ingredient for prediction of future returns for small leveraged equities, and we further use the predicted debt reductions for that purpose. Thus, the model favouring higher *sensitivity* over low *false positive rate* is desirable here. We can observe that out of the 35,276 observations predicted to reduce debt in the future, 20,019 really did so. So *Precision* is 57%. In terms of "No" outcomes, the detection rate (*specificity*) is 56%, meaning that 56% of the firms that actually failed to reduce their debt in the next twelve months were correctly predicted to fail.

Finally, we can conclude that we are able to predict the future ability to reduce long-term debt of a small leveraged high-value firm with accuracy of more than 62%. This means that the built predictor classifies over 62% of the cases correctly regardless of the actual outcome. Therefore, the probability of successfully classifying an observation is 8 percentage points higher than using simple unconditional probability (i.e., classifying all observations as "*No*"). Another overall performance measure for classification models is *Area under curve (AUC)*, where the curve is *receiver operating characteristic (ROC)* curve. ROC curve illustrates the model performance in terms of *sensitivity* and false positive rate for various probability thresholds. Figure 5.1 depicts ROC curve for our GBM model. Looking at the scale on the right-hand side and the color of the ROC curve, you can observe the model performance with respect to different probability thresholds. We can also observe that the ROC curve is smooth, and there is no specific threshold value at which the separability of the data would step up substantially.

Even though the choice of the applied classifier was made based on extensive academic research regarding similar classification problems (see Barboza *et al.*, 2017; Chingono & Rasmussen, 2016; Gepp *et al.*, 2010, etc.), we estimate logistic regression to provide a benchmark against a conventional regression method. The accuracy of the estimated logit model is 59%, which is lower than 62% for GBM. The GBM model scored 0.68 compared to 0.64 for the logit model in terms of AUC. The difference in performances of these two algorithms is narrower than suggested by a similar study of Chingono & Rasmussen (2016) though. In terms of the baseline model, GBM, we achieve similar performance to Chingono & Rasmussen (2016) with substantially more up-to-date and shorter in and out-of-sample data period.

Figure 5.1: ROC curve for Future debt paydown



The ROC curve depicts out-of-sample sensitivity (y-axis) and false positive rate (x-axis) for various probability thresholds. The dummy variable whether the company will be able to reduce its debt in the next twelve months is predicted using our GBM model over the 6.5-year-long out-of-sample period starting in April 2015. The individual probability thresholds are indicated by the color scale on the right-hand side of the figure.

Previous studies focused on debt or bankruptcy-related classification problems usually incorporated stock-level market data such as past returns in their classification models. We ignore such variables as the rationale behind their inclusion in those studies is not relevant here. Lagged returns can proxy for information not captured by fundamental data; however, this primarily applies to bankruptcy problems. Given the monthly frequency of our data, using past returns to predict debt paydown would create excessive noise. We considered the behavioral factors during the tuning phase of the model and they exhibited near-zero predictive power on the validation sample.

On the left-hand side of Figure B.1, we can observe relative importance of individual predictors. The two most important predictors are capital expenditures relative to assets and consensus estimate on growth rate of sales, followed by current ratio. The importance here means that introducing noise into these signals would increase the model's misclassification rate the most (see Section 4.3 for further detail regarding this methodology), e.g., losing the information contained in predictor CAPEX/Assets would decrease the model's accuracy by about 12.5 percentage points. It suggests that the future debt reduction mostly depends on whether companies invest in their fixed assets, i.e., those expanding property, plant, and equipment will rather take on more debt than deleverage. Such negative relationship is demonstrated in respective partial dependence plot in Figure B.2. The plots depict relationships between individual predictors and target variable while averaging away the effect of the others. Note that the relationship is not linear for none of the most important predictors. For example, the probability of paying down debt is increasing in *Sales growth estimate* only for negative growth rates. After that, the probability is decreasing with the lowest predicted probabilities as low as 30% for high-growth companies. This is consistent with the rationale that high-growing firms are likely rather to take on additional debt than deleverage.

In addition to the investment-related signals, liquidity and solvency ratios also ranked on top in terms of their predictive power, which is consistent with the universe composition, i.e. firms in financial distress are more likely to default on their debt obligations. EBITDA/EV valuation multiple is the least significant predictor. Therefore, it supports the rationale that only accounting-based variables are relevant for future debt repayment and suggests that consideration of market-based signals would be limited even without model supervision. Please see Appendix B to explore the predictive power of other signals and their marginal effects.

5.2 Mispricing strategy

Accurate prediction of future returns is the cornerstone in uncovering potential alpha returns. Adopting the machine-learning algorithm specified in Subsection 4.2.1, we are able to predict monthly returns with RMSE of 8.6 percentage points. With the standard deviation of the pooled leveraged value small-capitalization universe of 8.6 p.p. as well, the relative absolute error (RAE) is equal to 1. This relative level of accuracy is consistent across the entire testing period (i.e., it is approximately constant even when broken down into subperiods and calculated individually for each of them). The RAE is stable even for the calendar year 2020 despite increased volatility due to Covid-19 outbreak. R-squared is equal to 0. With respect to the observed model performance metrics (RMSE, RAE, R-squared, and correlation coefficient between predicted and true values), the gradient boosting machines model outperformed its OLS counterpart in all aspects, e.g., RMSE of 8.6 p.p. vs. 9.0. This is consistent with recent asset pricing literature supporting the application of machine-learning techniques over the conventional ones (see Gu *et al.*, 2020).

The right-hand side of Figure B.1 provides a visual representation of relative predictive power of the 25 most dominant signals. Cumulative total return over the last 6 months is clearly the most important one out of the 34 predictors considered by the GBM model. The estimated reduction in MSE attributable to the splits based on *6-month return (decimal)* is almost 6. Comparison to the MSE of fitted GBM model of 73[1] suggests the absolute importance of this predictor. The second most significant signal is *EBITDA/EV* with estimated MSE withholding of 2.9. This is not surprising as it serves as a proxy for value, one of the most prominent risk factors. Analogously, we can observe that also other well-documented risk factors such as profitability, investment, and size are represented by respective proxies among the most important variables. On top of these effects, many other anomalies such as intangibles, accruals, trading volume, or analyst estimates show considerable contribution. Though, the accuracy of the estimates of MSE reductions depends on the level of interactions

[1]Both estimated MSE reduction attributable to individual signals and the fitted GBM MSE value of 73 are based on in-sample tests.

between the signals.

Using *partial dependence plots (PDPs)* in Figure B.3 we can further elaborate on effects of the six most influential predictors. Each partial dependence plot shows the predicted future return by the GBM model for given predictor values while "averaging out"[2] the effect of the others. Visual inspection provides quick insight into the marginal effects across various values of the signals. Assuming that *ROA (CFO)* proxies for profitability, *Asset growth* proxies for firm's investment strategy and *EBITDA/EV* proxies for value, the estimated partial relationships correspond to those of (Fama & French, 2015) in terms of signs and importance.

The supremacy of *6-month return (decimal)* is consistent with past relevant literature since recent price actions are the most influential signals in the vast majority of recent predictive studies using monthly data (e.g. Gu *et al.*, 2020; Choi *et al.*, 2021, etc.). This past price action effect is usually in the form of momentum. In this thesis, we recognize momentum as well; however, this is true primarily for returns larger than 50% loss or lower than 25% gain in the last 6 months. For loss larger than 50%, the stocks tend rather to reverse as the marginal relationship is negative. Please see Figure B.3. Analogously for *EBITDA/EV*, we can assume that the future returns are increasing in value effect with flattening in the tails. Average monthly return in-sample[3] is about 1.0%. Thus, we can see for what values of the given predictor, the average future return prediction is above the in-sample average.

Before the mispricing strategy is established (i.e., we construct the long-short portfolio), we can investigate the relationship between predicted and realized returns, and the value added by the stock ranking. As described in Section 4.2, each month, all firms are sorted into deciles based on the predicted future returns. In Table 5.2 we can observe selected risk-adjusted performance indicators across the derived deciles for the out-of-sample period April 2015 - September 2021. Performance of the individual deciles relative to each other is another evaluation tool for the constructed GBM model.

[2]Please see Section 4.3

[3]Similar to variable importance testing, marginal effects are derived on in-sample data so *partial dependence plots* should be discussed with respect to in-sample data characteristics.

Table 5.2: Out-of-sample performance of decile portfolios

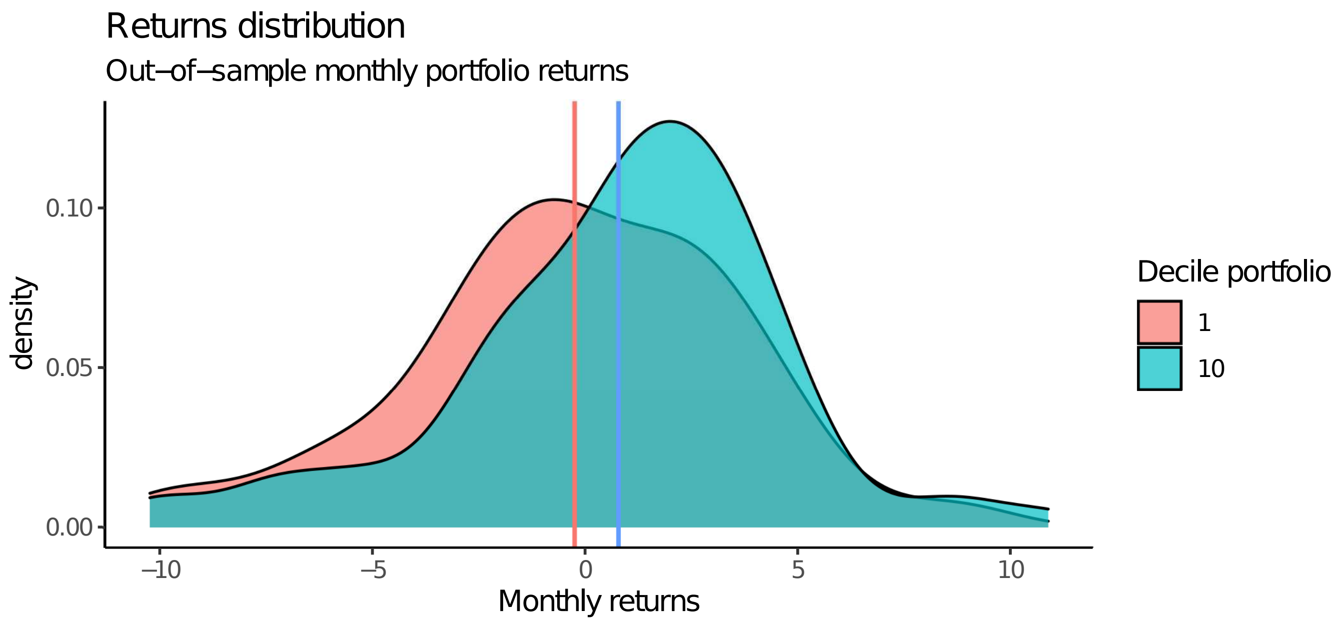
	1	2	3	4	5	6	7	8	9	10
Excess return (%)	-0.32	-0.05	0.44	0.48	0.58	0.36	0.69	0.66	0.74	0.71
SD (%)	3.84	3.76	3.62	3.25	3.54	3.50	3.75	3.88	3.94	3.87
Sharpe ratio	-0.08	-0.01	0.12	0.15	0.16	0.10	0.18	0.17	0.19	0.18
Alpha	-1.03	-0.78	-0.26	-0.16	-0.12	-0.33	-0.03	-0.09	-0.04	-0.03
Beta	0.78	0.79	0.76	0.69	0.75	0.75	0.79	0.81	0.84	0.81

The individual decile portfolios are constructed using sorts of predicted 1-month-ahead returns as described in Subsection 4.2.2. *Excess return (%)* denotes the average of monthly return on top of US Treasury bill rate over the out-of-sample period. *SD* denotes standard deviation of *Excess return (%)*. Sharpe ratio equals the excess return divided by standard deviation and represents the excess return per unit of risk. *Alpha* and *Beta* are based on CAPM regressions for each decile portfolio over the out-of-sample period from April 2015 to September 2021.

Monthly returns on top US Treasury bill rate are clearly increasing in the ranking proposed by our GBM model. The average monthly excess return of going long the bottom 10% of stocks with equal weighting and monthly rebalancing is -0.32%, whereas the most profitable long-only strategy would be to purchase stocks categorized in 9^{th} decile with average monthly return of 0.74%. Sharpe ratio provides a volatility-adjusted basis to asses returns and the pattern is obviously positive as well. The alpha and beta reported in Table 5.2 are based on CAPM[4] and the increasing values of alpha display the value added by the model. Even though the purpose of the table is to demonstrate the positive relationship between the portfolio sorts and their performance, the negative alpha return across the entire spectrum requires further discussion. The negativity is attributable to the universe we pick the stocks from - leveraged value small-capitalization stocks. Between 2015 and 2021, the portfolio tilting favoured growth over value and the premiums historically presented faded completely. The estimated alpha for the entire universe is about -0.12% indicating decent precision of the ranking by the model. During this out-sample period, the value premium measured by HML was -0.47% and the size premium SMB was slightly negative as well. Adjusting the positive exposure to such negative premiums (i.e., estimating the three-factor model by Fama & French (1993)) would reverse the negative alpha for the top 4 deciles into positive territory. Further factor exposure and abnormal returns analysis regarding top and bottom deciles is presented later in this section.

[4] With benchmark market portfolio for developed countries as defined by Kenneth R. French's data library.

Figure 5.2: Probability distribution functions of selected deciles



The decile portfolios are constructed using sorts of predicted 1-month-ahead returns as described in Subsection 4.2.2. The two vertical lines represent the average monthly returns for respective decile portfolios. Each probability distribution function is based on equal-weighted portfolio returns for 78 months in the out-of-sample period (April 2015 - September 2021).

Piotroski (2000) argues that the success of simple value investing strategy (based on book-to-market ratio) is driven by only a few outliers. Our mispricing strategy is robust in this matter due to removal of the outlying observations (in the *future returns* variable) for both model fitting and model testing. In Figure 5.2 we plot distribution of 78 out-of-sample monthly returns for 1^{st} and 10^{th} decile portfolios. The two vertical lines in the chart depict the means of the respective equal-weighted portfolios. Kurtosis of 3.22 for Decile 1 and 3.99 for Decile 10 suggest that the return series are not subject to extreme outliers. Moreover, both samples are left-skewed, suggesting that exceptional negative returns appear in both deciles. E.g., the worst-performing month was March 2020 due to uncertainty caused by global pandemic and both deciles exhibited almost the same returns in that month.

Each month, returns of the mispricing strategy are calculated as the average of return on stocks in Decile 10 and the return on selling short the stocks in Decile 1 with equal weighings and rebalancing every period. Please see Section 4.2 for an explicit description of how to invest in such a strategy. Since the mispricing hypothesis is primarily from a theoretical point of view, transaction costs are not considered when calculating the strategy returns. However, a discussion on transaction costs and the investability of the strategy will follow later in this section.

Table 5.3: Out-of-sample mispricing strategy testing

	Dependent variable: **Excess return**				
	Long-Short			Long leg	Short leg
	(1)	(2)	(3)	(4)	(5)
Alpha	0.416***	0.424**	0.337**	0.114	0.561**
	(0.120)	(0.124)	(0.112)	(0.189)	(0.195)
Market	0.015	0.012	-0.014	0.767***	-0.796***
	(0.028)	(0.029)	(0.029)	(0.050)	(0.051)
SMB		0.036	0.176*	0.572***	-0.219
		(0.083)	(0.080)	(0.136)	(0.140)
HML		0.009	0.154*	0.177	0.130
		(0.044)	(0.074)	(0.126)	(0.130)
CMA			-0.083	0.146	-0.313
			(0.132)	(0.224)	(0.231)
RMW			0.500***	0.205	0.796***
			(0.108)	(0.183)	(0.189)
Observations	77	77	77	77	77
R^2	0.004	0.007	0.240	0.849	0.834

The mispricing strategy is tested by regressing monthly excess returns of the long-short portfolio described in Section 4.2 on major risk factors (i.e., we use CAPM, and three and five-factor models by Fama & French). Variable *Market* denotes monthly total return for global market portfolio as defined by Kenneth R. French's data library. *SMB, HML, CMA* and *RMW* represent factor premiums (for small market capitalization, high book-to-market ratio, conservative investment policy and robust profitability; respectively) as defined by Fama & French. *Alpha* is the intercept and represents out-of-sample monthly abnormal risk-adjusted returns of the mispricing strategy, i.e., mispricing in leveraged value small-capitalization stocks.

. p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Table 5.3 provides table of the mispricing strategy tests using key asset pricing models, i.e., we regress excess returns of the long-short portfolio on several prominent risk factors. Adjusting for various risk factors (CAPM, FF3, and

FF5), the long-short portfolio yields significant alpha returns in the tested period starting May 2015. The machine-learning-based strategy is able to produce a return of 0.34% per month on top of the return that would be required by investors (by those investors who consider 5 risk factors - systematic risk, size risk, risk of financial distress, weak profitability, and aggressive investment policy). If adjusted only for risk proxied by market portfolio, the mispricing strategy yields 0.42 % risk-adjusted abnormal monthly return.

We can observe that the exposure to market risk premium is not statistically different from 0, which is consistent with the strategy that aims to exploit mispricing and is supposed to be market neutral, i.e., market performance does not affect the performance of this strategy. For individual legs of the mispricing strategy, the market exposure is significant and symmetric (see *(4)* and *(5)* in Table 5.3). The strategy slightly benefits from size premium, especially in its long leg. In the short leg, the coefficient is negative as anticipated but insignificant, which creates an overall tilt towards small-capitalization premium. Given that the average market capitalization is $ 617 million and $ 1.3 billion for stocks in the short leg and long leg, respectively, we can conclude that even though small size usually yield a risk premium relative to large size, such an effect does not hold within the small-capitalization spectrum. Otherwise, the exposure of the short leg would be significantly negative. In fact, partial dependence plot implies a rather positive relationship between size and return within the universe.

Another noteworthy finding is that both short and long components yield value premium even though we select the underlying stocks exclusively from high book-to-market region, i.e., the shorting is - according to past asset pricing literature - supposed to exhibit negative relation as this leg shorts the value stocks. Book-to-market ratio is about 1 for the stocks in the long leg and 1.4 for those in the short one. Moreover, partial dependence plot showed that our algorithm suggests a negative relationship between book-to-market valuation and future return within the universe of leveraged value small-caps, especially for stocks with *book-to-market ratio* over 0.8. While prior studies document the B/M risk factor well, the findings are usually based on a broad sample of stocks. Based on the significant coefficient of 0.154 for HML, we argue that this value effect is not linear across the whole spectrum of B/M values and turns negative for stocks with too high B/M multiples.

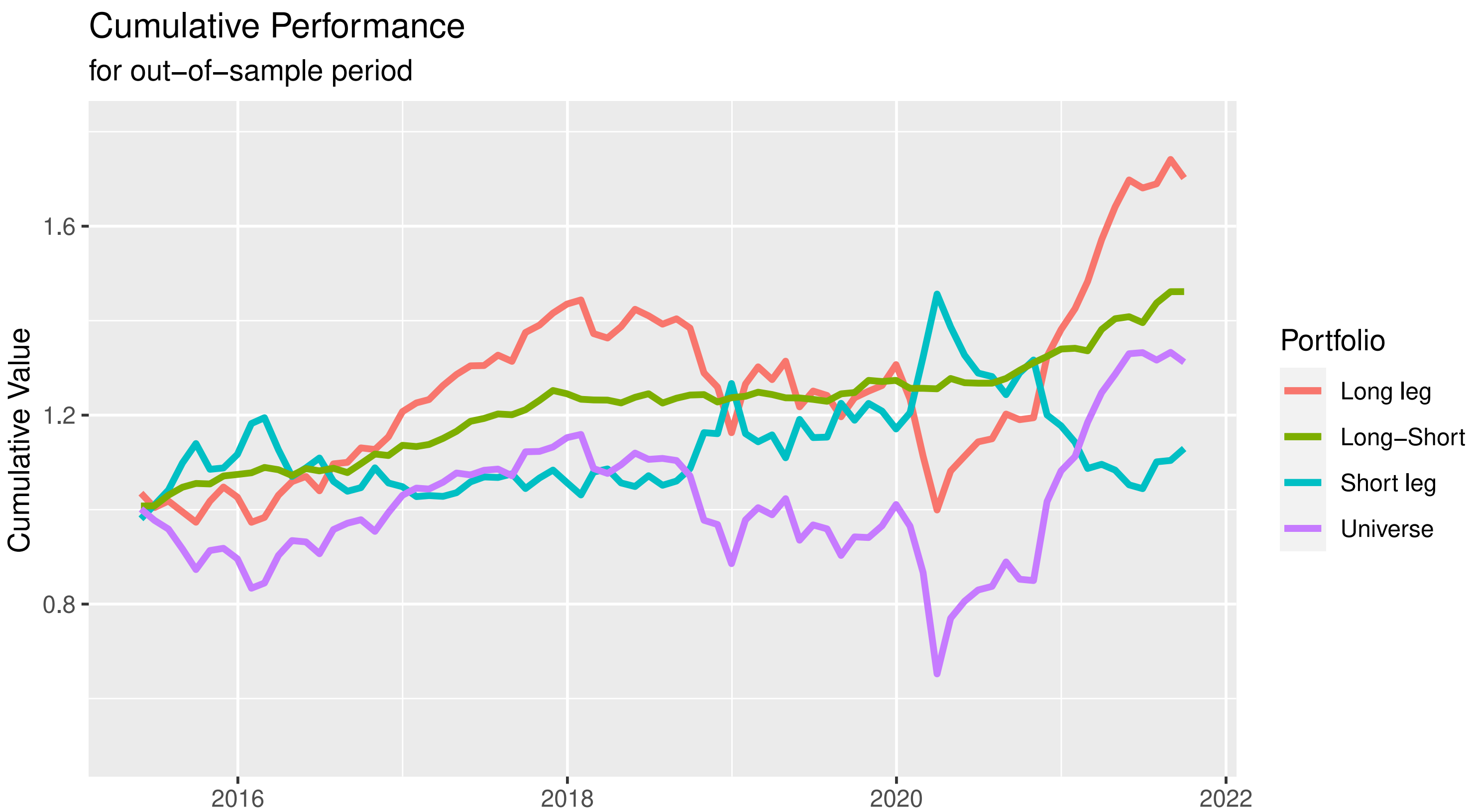
The relationship with investment policy is represented by CMA factor in the table. The mispricing strategy is neutral to this factor since the performance of the invested stocks is not affected in either leg. This is slightly surprising since Asset growth - which serves as a proxy for investment policy factor in several prior studies - is the fourth most important predictor of our GBM model, with the marginal effect being clearly negative.

Conversely, the strategy is substantially loaded in profitability premium. The universe from which the algorithm selects the stocks is not restricted to any profitability values. Hence the algorithm can exploit the profitability premium across the entire spectrum of profitability values (unlike for value factor where the potential to capitalize on that premium in full is limited by definition of the universe). The average operating profitability [5] is monotonically increasing in the portfolio ranking with average operating profitability of -12.5% for Decile 1 and 18% for Decile 10. Therefore, statistically positive exposure in the short leg of the strategy indicates that selling short stocks with weak profitability enhances expected excess return by 0.8 percentage points per month. Coefficient for long leg is positive favouring long position into companies with robust profitability; however, the statistical significance is rather low. Overall the mispricing strategy yields a positive profitability beta return.

The monthly alpha for the short component of the mispricing strategy is 0.56 percentage points after controlling for key risk factors. On the other hand, the long component fails to reject the zero-alpha hypothesis. Apparently, the short leg in the mispricing strategy is the primary driver of the abnormal returns. Contrariwise in terms of the returns not adjusted for various sources of risk, long positions outperform the short ones. Please see Figure 5.3 for cumulative performance based on the mispricing strategy, the long-only component, the short-only component, and the equal-weighted portfolio of the leveraged value small-capitalization stocks (=the universe). The mispricing strategy is superior to simply holding the universe over the entire observed period. The volatility of the strategy is also considerably lower.

[5] Operating profitability in this context is defined according to Fama & French (2015) as operating income less interest expense divided by book equity.

Figure 5.3: Mispricing strategy cumulative performance



The figure depicts out-of-sample cumulative performance of the mispricing strategy (i.e. machine-learning-based long-short portfolio as shown in Table 5.3 and defined in Subsection 4.2.2), its long-only, and short-only components; and equal-weighted portfolio of leveraged value small-capitalization stocks (i.e., the universe defined in Section 3.2).

Universe of leveraged value small-caps likely suffers from lack of coverage by analysts due to small size and elevated risk of financial distress reflected by attractive valuation multiples. Such characteristics might represent an ideal environment for mispricing by investors to take place, especially in periods of elevated market volatility or credit crunches. If investors are really more susceptible to misprice in these periods, the constructed strategy will be able to capitalize on that, i.e., yield higher abnormal returns than usual. In Table 5.4 we present results of regressing Jensen's alpha - which represents the magnitude of the abnormal return each month (see Equation 4.5 for greater detail) - on option-adjusted spread on high-yield bonds *(1)*, and average monthly value of index volatility VIX *(2)*. We can see that neither of the two hypothesized relationships is substantiated by the data, i.e., the coefficient is not statistically positive in either case. Thus, we found no evidence that the mispricing strategy performs (on a risk-adjusted basis) better during periods of lifted volatility or tighter credit conditions.

Table 5.4: Mispricing during high volatility and tight credit

	Dependent variable:	
	Jensen's alpha	
	(1)	(2)
Intercept	0.348	0.304
	(0.432)	(0.307)
OAS	0.015	
	(0.091)	
VIX price		0.006
		(0.016)
Observations	77	77
R^2	0.0004	0.002
F Statistic (df = 1; 75)	0.027	0.157

Jensen's alpha represents the magnitude of abnormal returns yielded by mispricing strategy (i.e. yielded by machine-learning-based long-short portfolio as shown in Table 5.3 and defined in Subsection 4.2.2) in each particular month after adjusting for market risk of CAPM. *OAS* is option-adjusted spread between high-yield US bonds and US spot treasury curve. *VIX price* is a measure of expected volatility at S&P 500 in the next 30 days as implied by option prices. Null hypothesis is that mispricing does not change with different credit conditions, or market volatility, *(1)* and *(2)* respectively.

. p<0.1; * p<0.05; ** p<0.01; *** p<0.001

Mispricing strategy beats the market on a risk-adjusted basis (see Table 5.3). Put adjustments for various beta factors aside, the cumulative total return on the global market portfolio is 92% as opposed to the return to mispricing strategy of 46%. So the market performed better on absolute basis. However, it does not necessarily imply that the superior strategy here is to hold the market portfolio instead. According to portfolio selection theory (see Markowitz, 1952; Tobin, 1958), the investors should invest in such a risky portfolio that the Sharpe ratio is maximized, and incorporate a position in risk-free asset to achieve the absolute return of personal preference. Over the observed out-of-sample period, Sharpe ratio[6] is equal to 0.42 for the long-short portfolio and

[6] Please note the Sharpe ratio values are calculated using monthly data.

0.21 for global market portfolio[7]. Though adopting exclusively the mispricing strategy is not optimal according to portfolio selection theory. Considering the imperfect correlation of these two portfolios, the optimal risky portfolio would consist of 91% of the long-short mispricing strategy-based portfolio and 9% of the global market portfolio due to diversification.

So far, the analysis does not take into account any transaction costs. Given that the short leg is the main contributor to the positive alpha of the strategy and the small-capitalization nature of the universe, transaction costs might not be necessarily negligible in this particular region of the portfolio. Average market capitalization of the stocks sold short within this leg is USD 800 mil. Using a proprietary database of a large financial institution on security borrowing fees[8], Bekjarovski (2018) reports an average borrowing rate for similar-sized US companies of 2.4% p.a. Adopting this finding as an assumption here, it would represent 35% of the short leg monthly alpha (see (5) in Table 5.3). Manual check[9] on borrowing rates for several micro-capitalization stocks included in our portfolio indicated rather lower than the assumed cost of borrowing. Through there is potential for outliers in case of a deficient number of shares available for borrowing. While this poses a noteworthy restriction to executing the strategy in full, it is still investible given that on an average month, the short leg consists of 90 stocks.

[7]Global market portfolio for developed countries by Kennetch R. French data library.

[8]In order to short sell a security, it needs to be borrowed. For borrowing a security, an annualized borrowing rate is charged.

[9]Such manual check was carried out at the end of 2021 using borrowing rates quoted by a global brokerage.

6 Conclusion

The focal point of this thesis is the universe of leveraged values small-capitalization stocks, which are characterized by their exposure to common risk factors. We show that the characteristic environment of this universe is also associated with mispricing of the securities by investors. We also identify the major anomalies that contribute to the predictability of returns and drive the abnormal returns of the mispricing strategy.

First, we predict the ability of the companies to reduce long-term debt in the course of the next twelve months as - according to prior literature - this might be the primary driver of abnormal returns in such a universe. Having gradient boosting machines algorithm learning on 15 years of fundamental-only stock-level monthly data, we are able to predict future deleveraging with out-of-sample accuracy of 62%. The most important predictor in this matter is *CAPEX/Assets* implying that higher CAPEX reduces the probability of deleveraging in the near future, though not linearly. The other predictors that reduced the out-of-sample misclassification rate the most are consensus estimate of growth rate of sales (negative relationship for growing firms) and current ratio (positively related to the probability of deleveraging). Even though the predictions are ahead as much as 12 months, we are able to take advantage of the monthly frequency of data due to LTM accounting figures. As a result, most recent accounting data is considered, e.g., every quarterly release is immediately reflected. Even though the boosted trees algorithm is often touted as a "black-box" technique due to its limited interpretability, we present average partial prediction functions for individual predictors and estimates of their predictive power allowing for further detail into the deleveraging determinants (see Appendix B).

Subsequently, we formulate the mispricing strategy by sorting the universe into deciles based on predicted future returns and by going long the top decile and selling short the bottom one. We predicted 1-month ahead returns over 6.5-years-long out-of-sample period starting in April 2015 with RMSE of 8.6 (i.e., equal to one standard deviation). The average excess return is -0.32% for the Decile 1 portfolio and 0.72% for Decile 10 portfolio. The decile portfolios are clearly increasing in Sharpe ratios and alphas determined by CAPM, validating the stock selecting process.

Finally, the constructed long-short portfolio yields abnormal risk-adjusted returns of at least 0.34 percentage points per month according to all three asset pricing models used for testing (CAPM, and three and five-factors models of Fama & French). Since empirical asset pricing is a saturated research area, several differential aspects should not be left unnoticed. We exclusively focus on leveraged value small-capitalization universe, allowing for relationships specific to companies with such characteristics. In addition, we apply boosted trees methodology that can capture complex relationships between the predictors and the future returns. We also present the predictive contribution of individual signals and their nonlinear partial effects. Our monthly stock-level data is aligned in terms of announcement dates and include delisted securities eliminating *look-ahead* and *survivorship bias*. Thus, on each portfolio rebalancing date, the algorithm considers the most recent and only already-announced information, mimicking a real information set an investor would have at his disposal.

Sharpe ratio of the constructed strategy is 0.42, exceeding the global benchmark. Based on mutual correlation, the most efficient portfolio (considering only these two assets) is composed of 91% mispricing strategy and 9% global benchmark portfolio. However, security borrowing costs associated with the short leg of the strategy would consume approximately 35% of the monthly alpha return attributable to that leg if borrowing rate of 2.4% p.a. is assumed. Since security borrowing rates are a function of lending supply and borrowing demand, excessive rates might eventually occur. Detailed analysis of borrowing rates is beyond the scope of this thesis. Further research on this matter would help validate the mispricing strategy's convertibility into real-world profits. In spite of potentially further need for clarity on the self-sufficiency of the strategy, it can readily be used in practice for stock screening. Another potential application is determination of a proper discount rate, especially for short-term periods.

Compared to previous studies addressing similar research questions, our algorithms were trained and tested on a considerably shorter and more recent period (from March 2000 to September 2021). While this allows to capture investors' behavior and asset pricing trends of the current millennium within the mispricing strategy, fundamental-only oriented model predicting future debt paydown would likely benefit from a more extended training period. Choice of the suitable machine-learning algorithm and associated specifications was based on several past papers providing relevant benchmarks (e.g. Gu *et al.*, 2020). Though, further potential in terms of predictive power might lie in alternative model specifications such as considering additional hyperparameters or other loss functions, e.g., using Huber loss function[1] poses an alternative treatment of outlying observations.

[1]Huber loss function defined by Huber (1964) combines squared-error L_2 and absolute error L_1 loss functions such that it switches from L_2 to L_1 for residuals greater than a specified maximum error value δ.

Bibliography

AMIHUD, Y. (2002): "Illiquidity and stock returns: cross-section and time-series effects." *Journal of Financial Markets* **5(1)**: pp. 31–56.

AMIHUD, Y. & H. MENDELSON (1986): "Asset pricing and the bid-ask spread." *Journal of Financial Economics* **17(2)**: pp. 223–249.

ANNAERT, J., J. CROMBEZ, B. SPINEL, & F. VAN HOLLE (2002): "Value and size effect: Now you see it, now you don't."

BANZ, R. W. (1981): "The relationship between return and market value of common stocks." *Journal of Financial Economics* **9(1)**: pp. 3–18.

BARBOZA, F., H. KIMURA, & E. ALTMAN (2017): "Machine learning models and bankruptcy prediction." *Expert Systems with Applications* **83**: pp. 405–417.

BASU, S. (1983): "The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence." *Journal of Financial Economics* **12(1)**: pp. 129–156.

BEKJAROVSKI, F. (2018): "How Do Short Selling Costs and Restrictions Affect the Profitability of Stock Anomalies?" *SSRN Electronic Journal* .

BELO, F., V. GALA, J. SALOMAO, & M. A. VITORINO (2019): "Decomposing firm value." *Technical report*, National Bureau of Economic Research.

BHANDARI, L. C. (1988): "Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence." *The Journal of Finance* **43(2)**: pp. 507–528.

BONDT, W. F. M. D. & R. THALER (1985): "Does the Stock Market Overreact?" *The Journal of Finance* **40(3)**: pp. 793–805.

Bradshaw, M. T., S. A. Richardson, & R. G. Sloan (2006): "The relation between corporate financing activities, analysts' forecasts and stock returns." *Journal of Accounting and Economics* **42(1)**: pp. 53–85.

Breiman, L. (2001): "Random Forests." *Machine Learning 2001 45:1* **45(1)**: pp. 5–32.

Brown, S. J., W. N. Goetzmann, & S. A. Ross (1995): "Survival." *The Journal of Finance* **50(3)**: pp. 853–873.

Chan, K. C. & N.-f. Chen (1991): "Structural and Return Characteristics of Small and Large Firms." *The Journal of Finance* **46(4)**: pp. 1467–1484.

Chan, L. K., Y. Hamao, & J. Lakonishok (1991): "Fundamentals and Stock Returns in Japan." *The Journal of Finance* **46(5)**: pp. 1739–1764.

Chan, L. K. C., J. Lakonishok, & T. Sougiannis (2001): "The Stock Market Valuation of Research and Development Expenditures." *The Journal of Finance* **56(6)**: pp. 2431–2456.

Chingono, B. K. & D. Rasmussen (2016): "Forecasting Debt Paydown Among Leveraged Equities: Analysis of US Stocks from 1964-2012." *SSRN Electronic Journal* .

Choi, D., W. Jiang, & C. Zhang (2021): "Alpha Go Everywhere: Machine Learning and International Stock Returns." *SSRN Electronic Journal* .

Chopra, N., J. Lakonishok, & J. R. Ritter (1992): "Measuring abnormal performance: Do stocks overreact?" *Journal of Financial Economics* **31(2)**: pp. 235–268.

Cohen, D. A. & P. Zarowin (2010): "Accrual-based and real earnings management activities around seasoned equity offerings." *Journal of Accounting and Economics* **50(1)**: pp. 2–19.

Constantinides, G. (1983): "Capital Market Equilibrium with Personal Tax." *Econometrica* **51(3)**: pp. 611–636.

Cook, T. J. & M. S. Rozeff (1984): "Size and earnings/price ratio anomalies: One effect or two?" *Financial Review* **19(3)**: p. 32.

Cooper, M. J., H. Gulen, & M. J. Schill (2008): “Asset Growth and the Cross-Section of Stock Returns.” *The Journal of Finance* **63(4)**: pp. 1609–1651.

Daniel, G., D. Sornette, & P. Wohrmann (2008): “Look-ahead benchmark bias in portfolio performance evaluation.”

Davis, J. L. (1994): “The Cross-Section of Realized Stock Returns: The Pre-COMPUSTAT Evidence.” *The Journal of Finance* **49(5)**: pp. 1579–1593.

Dhaliwal, D., S. Heitzman, & O. Zhen Li (2006): “Taxes, Leverage, and the Cost of Equity Capital.” *Journal of Accounting Research* **44(4)**: pp. 691–723.

Eisfeldt, A. L., E. Kim, & D. Papanikolaou (2020): “Intangible Value.”

Eisfeldt, A. L. & D. Papanikolaou (2013): “Organization Capital and the Cross-Section of Expected Returns.” *The Journal of Finance* **68(4)**: pp. 1365–1406.

Fairfield, P. M., J. S. Whisenant, & T. L. Yohn (2003): “Accrued Earnings and Growth: Implications for Future Profitability and Market Mispricing.” *The Accounting Review* **78(1)**: pp. 353–371.

Fama, E. F. & K. R. French (1992): “The Cross-Section of Expected Stock Returns.” *The Journal of Finance* **47(2)**: pp. 427–465.

Fama, E. F. & K. R. French (1993): “Common risk factors in the returns on stocks and bonds.” *Journal of Financial Economics* **33(1)**: pp. 3–56.

Fama, E. F. & K. R. French (2008): “Dissecting Anomalies.” *The Journal of Finance* **63(4)**: pp. 1653–1678.

Fama, E. F. & K. R. French (2015): “A five-factor asset pricing model.” *Journal of Financial Economics* **116(1)**: pp. 1–22.

Fama, E. F. & J. D. MacBeth (1973): “Risk, Return, and Equilibrium: Empirical Tests.” *Journal of Political Economy* **81(3)**: pp. 607–636.

Friedman, J. H. (2001): “Greedy Function Approximation: A Gradient Boosting Machine.” *The Annals of Statistics* **29(5)**: pp. 1189–1232.

GEORGE, T. J. & C.-Y. HWANG (2004): "The 52-Week High and Momentum Investing." *The Journal of Finance* **59(5)**: pp. 2145–2176.

GEPP, A., K. KUMAR, & S. BHATTACHARYA (2010): "Business failure prediction using decision trees." *Journal of Forecasting* **29(6)**: pp. 536–555.

GIBBONS, M. R., S. A. ROSS, & J. SHANKEN (1989): "A Test of the Efficiency of a Given Portfolio." *Econometrica* **57(5)**: pp. 1121–1152.

GRAHAM, B., D. DODD, & S. COTTLE (1962): *Security Analysis: Principles and Techniques*. McGraw-Hill.

GREEN, J., J. R. M. HAND, & X. F. ZHANG (2013): "The supraview of return predictive signals." *Review of Accounting Studies* **18(3)**: pp. 692–730.

GREENWELL, B. M. (2017): "pdp: An R package for constructing partial dependence plots." *R Journal* **9(1)**: pp. 421–436.

GU, S., B. KELLY, & D. XIU (2020): "Empirical Asset Pricing via Machine Learning." *The Review of Financial Studies* **33(5)**: pp. 2223–2273.

HAUGEN, R. A. (1995): *The new finance : the case against efficient markets*. Englewood Cliffs (N.J.) : Prentice-Hall.

HAUGEN, R. A. & N. L. BAKER (1996): "Commonality in the determinants of expected stock returns." *Journal of Financial Economics* **41(3)**: pp. 401–439.

HOU, K., C. XUE, & L. ZHANG (2020): "Replicating Anomalies." *The Review of Financial Studies* **33(5)**: pp. 2019–2133.

HUBER, P. J. (1964): "Robust Estimation of a Location Parameter." *The Annals of Mathematical Statistics* **35(1)**: pp. 73–101.

HWANG, S. & S. E. SATCHELL (1999): "Modelling emerging market risk premia using higher moments." *International Journal of Finance and Economics* **4(4)**: pp. 271–296.

IKENBERRY, D., J. LAKONISHOK, & T. VERMAELEN (1995): "Market underreaction to open market share repurchases." *Journal of Financial Economics* **39(2)**: pp. 181–208.

JACOBS, H. & S. MÜLLER (2018): "...And Nothing Else Matters? On the Dimensionality and Predictability of International Stock Returns." *SSRN Electronic Journal* .

JAFFE, J., D. B. KEIM, & R. WESTERFIELD (1989): "Earnings Yields, Market Values, and Stock Returns." *The Journal of Finance* **44(1)**: pp. 135–148.

JEGADEESH, N. (1990): "Evidence of Predictable Behavior of Security Returns." *The Journal of Finance* **45(3)**: pp. 881–898.

JEGADEESH, N. & S. TITMAN (1993): "Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency." *The Journal of Finance* **48(1)**: pp. 65–91.

JENSEN, M. C. (1968): "The Performance of Mutual Funds in the Period 1945-1964." *The Journal of Finance* **23(2)**: p. 389.

JENSEN, M. C., F. BLACK, & M. S. SCHOLES (1972): "The capital asset pricing model: Some empirical tests." *Studies in the theory of capital markets, Praeger Publishers Inc.* .

KEIM, D. B. (1983): "Size-related anomalies and stock return seasonality. Further empirical evidence." *Journal of Financial Economics* **12(1)**: pp. 13–32.

KIM, T. H. & H. WHITE (2004): "On more robust estimation of skewness and kurtosis." *Finance Research Letters* **1(1)**: pp. 56–73.

KNEZ, P. J. & M. J. READY (1997): "On the Robustness of Size and Book-to-Market in Cross-Sectional Regressions." *The Journal of Finance* **52(4)**: pp. 1355–1382.

KOT, H. W. & K. CHAN (2006): "Can contrarian strategies improve momentum profits." *Journal of Investment Management* **4(1)**.

KOTHARI, S. P., J. SHANKEN, & R. G. SLOAN (1995): "Another Look at the Cross-Section of Expected Stock Returns." *The Journal of Finance* **50(1)**: pp. 185–224.

LAKONISHOK, J., A. SHLEIFER, & R. W. VISHNY (1994): "Contrarian Investment, Extrapolation, and Risk." *The Journal of Finance* **49(5)**: pp. 1541–1578.

LEE, C. M. C. & B. SWAMINATHAN (2000): "Price Momentum and Trading Volume." *The Journal of Finance* **55(5)**: pp. 2017–2069.

LEIRVIK, T., S. R. FISKERSTRAND, & A. B. FJELLVIKÅS (2017): "Market liquidity and stock returns in the Norwegian stock market." *Finance Research Letters* **21**: pp. 272–276.

LEUNG, E., H. LOHRE, D. MISCHLICH, Y. SHEA, & M. STROH (2021): "The Promises and Pitfalls of Machine Learning for Predicting Stock Returns." *SSRN Electronic Journal* .

LEWELLEN, J. (2015): "The Cross-section of Expected Stock Returns." *Critical Finance Review* **4(1)**: pp. 1–44.

LINTNER, J. (1965): "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets." *The Review of Economics and Statistics* **47(1)**: p. 13.

LOUGHRAN, T. & J. RITTER (1995): "The New Issues Puzzle." *The Journal of Finance* **50(1)**: pp. 23–51.

MARKOWITZ, H. (1952): "Portfolio Selection." *The Journal of Finance* **7(1)**: pp. 77–91.

MCLEAN, R. D. & J. PONTIFF (2016): "Does Academic Research Destroy Stock Return Predictability?" *The Journal of Finance* **71(1)**: pp. 5–32.

MOSKOWITZ, T. J. & M. GRINBLATT (1999): "Do Industries Explain Momentum?" *The Journal of Finance* **54(4)**: pp. 1249–1290.

NOVY-MARX, R. (2013): "The other side of value: The gross profitability premium." *Journal of Financial Economics* **108(1)**: pp. 1–28.

PIOTROSKI, J. D. (2000): "Value Investing: The Use of Historical Financial Statement Information to Separate Winners from Losers." *Journal of Accounting Research* **38**: p. 1.

POTERBA, J. M. & S. J. WEISBENNER (2001): "Capital Gains Tax Rules, Tax-loss Trading, and Turn-of-the-year Returns." *The Journal of Finance* **56(1)**: pp. 353–368.

RASMUSSEN, D. & B. K. CHINGONO (2015): "Leveraged Small Value Equities." *SSRN Electronic Journal* .

Reinganum, M. R. (1981): "Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values." *Journal of Financial Economics* **9(1)**: pp. 19–46.

Richardson, S. A., R. G. Sloan, M. T. Soliman, & Tuna (2006): "The Implications of Accounting Distortions and Growth for Accruals and Profitability." *The Accounting Review* **81(3)**: pp. 713–743.

Ritter, J. (1991): "The Long-Run Performance of initial Public Offerings." *The Journal of Finance* **46(1)**: pp. 3–27.

Rosenberg, B., K. Reid, & R. Lanstein (1985): "Persuasive evidence of market inefficiency." *The Journal of Portfolio Management* **11(3)**: pp. 9–16.

Sharpe, W. F. (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk*." *The Journal of Finance* **19(3)**: pp. 425–442.

Sharpe, W. F. (1966): "Mutual Fund Performance." *The Journal of Business* **39(1)**: pp. 119–138.

Sloan, R. G. (1996): "Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings?" *The Accounting Review* **71(3)**: pp. 289–315.

Spiess, D. & J. Affleck-Graves (1995): "Underperformance in long-run stock returns following seasoned equity offerings." *Journal of Financial Economics* **38(3)**: pp. 243–267.

Thaler, R. H. (1987): "Anomalies: The January Effect." *Journal of Economic Perspectives* **1(1)**: pp. 197–201.

Titman, S., K. C. J. Wei, & F. Xie (2004): "Capital Investments and Stock Returns." *The Journal of Financial and Quantitative Analysis* **39(4)**: pp. 677–700.

Tobek, O. & M. Hronec (2021): "Does it pay to follow anomalies research? Machine learning approach with international evidence." *Journal of Financial Markets* **56**: p. 100588.

Tobin, J. (1958): "Liquidity Preference as Behavior Towards Risk." *The Review of Economic Studies* **25(2)**: pp. 65–86.

Tukey, J. W. (1977): *Exploratory data analysis*. Reading, Mass.: Addison-Wesley Pub. Co.

Vassalou, M. & Y. Xing (2004): "Default Risk in Equity Returns." *The Journal of Finance* **59(2)**: pp. 831–868.

Youden, W. J. (1950): "Index for rating diagnostic tests." *Cancer* **3(1)**: pp. 32–35.

Zhang, H. (2005): "Share price performance following actual share repurchases." *Journal of Banking & Finance* **29(7)**: pp. 1887–1901.

[LO]. Conclusion

A Appendix A: Variable Definitions

Table A.1: List of variables

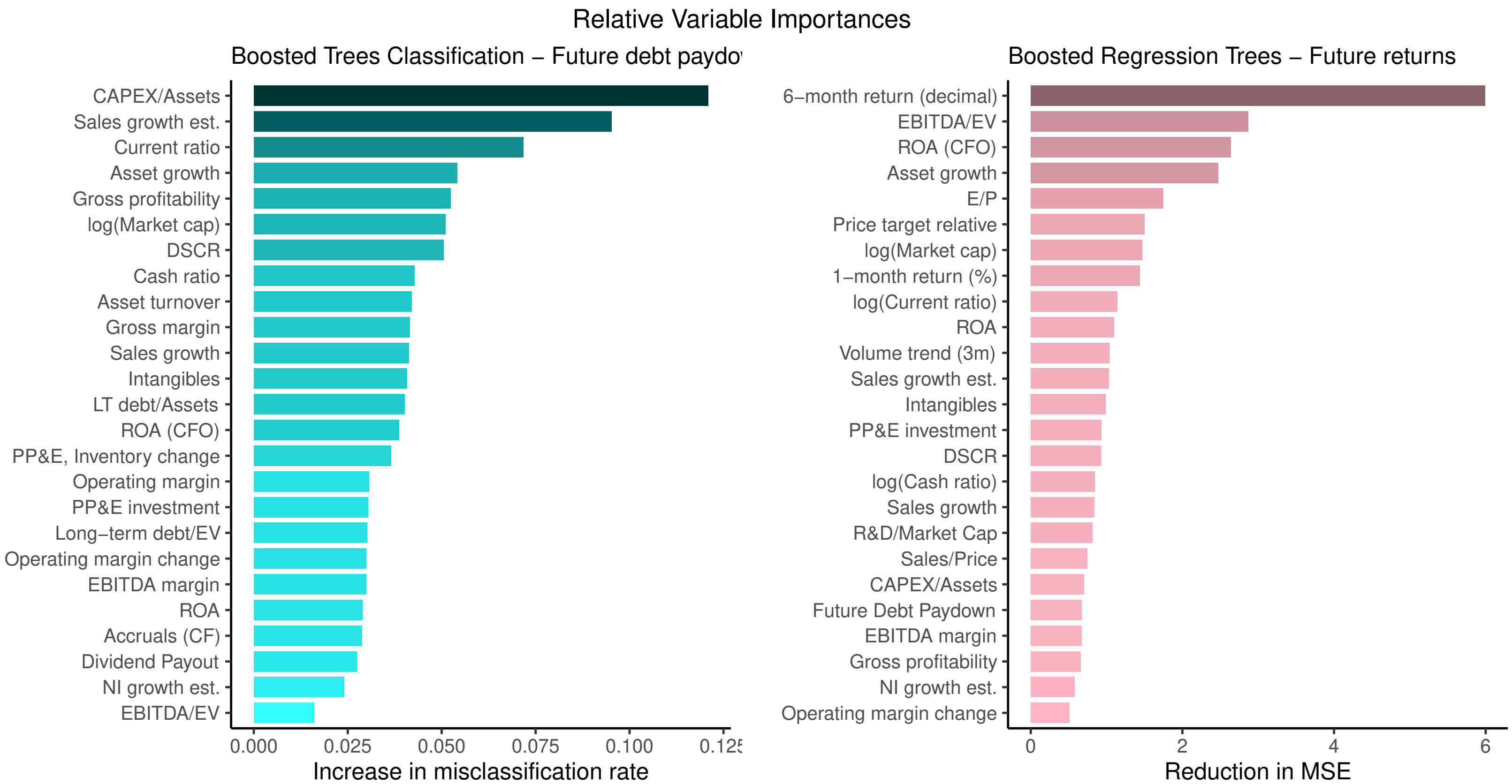
Variable	Description	Debt paydown model	Total return model
EBITDA/EV	EBITDA / Enterprise value	✓	✓
LT debt/EV	Long-term debt / Enterprise value	✓	✓
LT debt/Assets	Long-term debt / Total assets	✓	
log(Market cap)	Natural logarithm of market capitalization	✓	✓
DSCR	Operating income / Debt service	✓	✓
Current ratio	Current asset / Current liabilities	✓	
Cash ratio	Cash and short-term investments / Current liabilities	✓	
log(Current ratio)	Natural logarithm of Current ratio		✓
log(Cash ratio)	Natural logarithm of Cash ratio		✓
ROA	Net Income / Total assets	✓	✓
ROA (CFO)	Operating CF / Total assets	✓	✓
Gross margin	Gross profit / Total revenue	✓	✓
Gross profitability	Gross profit / Total assets	✓	✓
EBITDa margin	EBITDA / Total revenue	✓	✓
Operating margin	Operating income / Total revenue	✓	
Operating margin change	Operating margin$_t$ − Operating margin$_{t-12}$	✓	✓
Accruals (CF)	(Net income - Operating CF) / Total assets	✓	✓

PP&E investment	(PP&E / Total assets)$_t$ − (PP&E / Total assets)$_{t-12}$	✓	✓
PP&E and inventory change	(PP&E and inventory/ Total assets)$_t$ − (PP&E and inventory/ Total assets)$_{t-12}$	✓	
Intangibles	Intangibles / Total assets	✓	✓
Dividend Payout	Cash dividends / Net income	✓	✓
Sales growth	Total revenue$_t$ / Total revenue$_{t-12} - 1$	✓	✓
Asset growth	Total assets$_t$ / Total assets$_{t-12} - 1$	✓	✓
CAPEX/Assets	CAPEX / Total assets	✓	✓
CAPEX/Assets dummy$_t$	(CAPEX/Total assets)$_t$ > (CAPEX/Total assets)$_{t-12}$		✓
NI growth est.	Consensus estimate of net income growth rate over next twelve months	✓	✓
Sales growth est.	Consensus estimate of total sales growth rate over next twelve months	✓	✓
1-month return (%)	Total return over the last month, including dividends. Expressed in percentage.		✓
6-month return (decimal)	Total return over the last six months, including dividends. Expressed in decimals.		✓
Debt Paydown$_t$	=1 if *Long-term debt*$_t$ > *Long-term debt*$_{t-12}$, and 0 otherwise		✓
Future Debt Paydown$_t$	Estimated probability of *Debt Paydown*$_{t+12} = 1$, i.e. probability of reducing long-term debt in next twelve months. This estimate represents the target variable for the Debt paydown model.		✓
Volume trend (3m)	Average daily volume$_t$ / Average daily volume$_{t-3}$		✓
Price target relative	Consensus share price target for next twelve months / Current share price. Adjusted for stock splits.		✓
R&D/Market Cap	Research and development expense / Market capitalization		✓
Extra items % NI	Net income before extra items / Net income -1		✓

Dividend Yield	Cash dividend per share / Share price. Adjusted for stocks splits.		✓
Sales/Price	Total revenue / Market capitalization		✓
E/P	Net income / Market capitalization		✓
Book-to-Market	Book value of equity / Market capitalization		✓
Asset turnover	Total revenue / Total assets	✓	

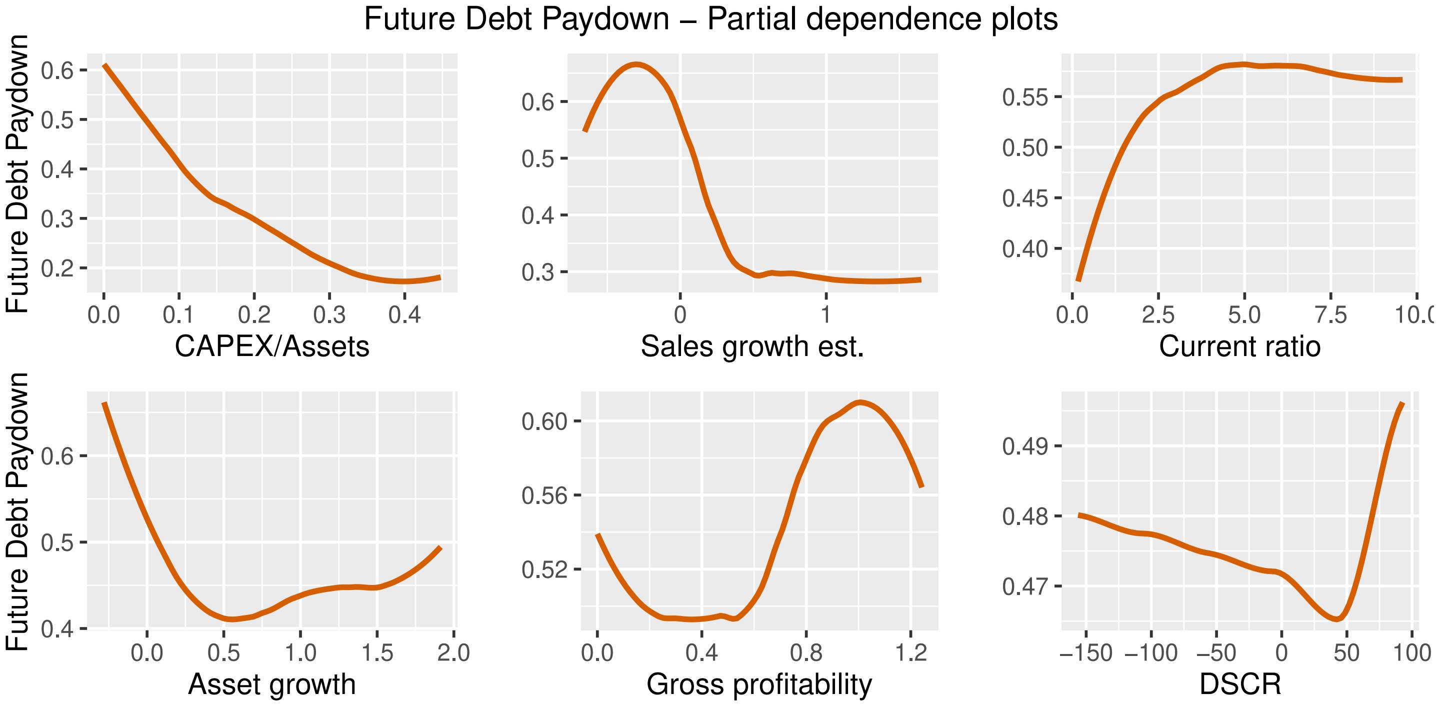
B Appendix B: Model Interpretation

Figure B.1: Variable importances in the predictive models



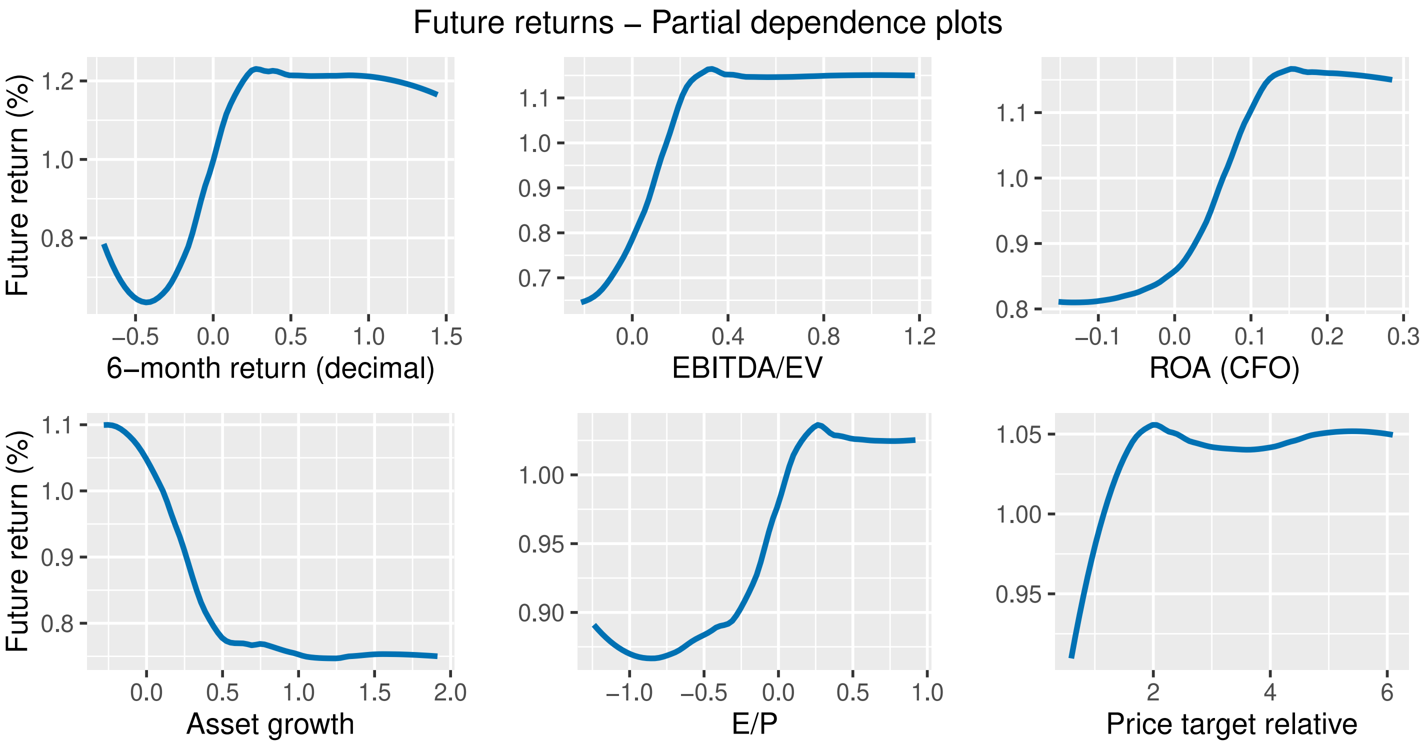
Left-hand side of the figure shows what would be the increase in missclasification rate of our future debt paydown prediction model if we "noised up" the individual signals while keeping the others unchanged. The right-hand side of the figure shows the reduction of mean squared errors for predicting future monthly returns attributable to the individual predictors. In both cases the underlying predictive model is gradient boosting machines with trees. Both the model fitting and the metrics are based on universe of leveraged value small-capitalization stocks for in-sample period (March 2000 - February 2015). Detailed calculations are provided in Section 4.3.

Figure B.2: Future Debt Paydown - Partial dependency



The figure depicts estimated relationships between the probability of reducing long-term debt in next welve months and selected individual predictors while averaging out the effects of the other predictors. The underlying predictive model is gradient boosting machines that considers 25 lagged fundamental signals and is trained on leveraged value small-capitalization stocks between March 2000 and February 2015. See detailed methodology in Section 4.3

Figure B.3: Future Return - Partial dependency



The figure depicts estimated relationships between 1-month ahead return and selected individual predictors while averaging out the effects of the other predictors. The underlying predictive model is gradient boosting machines that considers 34 lagged stock-level signals and is trained on leveraged value small-capitalization stocks between March 2000 and February 2015. See detailed methodology in Section 4.3.