Aalto University
School of Business

DISTRESS ANOMALY REVISITED

Distance to default in the US stock market

Master's Thesis
Jan (Kristian) Koskinen
Aalto University School of Business
Master's Programme in Economics
Fall 2022

Author Jan (Kristian) Koskinen

Title of thesis Distress anomaly revisited – Distance to default in the US stock market

Degree Master of Science in Economics and Business Administration

Degree programme Master's Programme in Economics

Thesis advisor(s) Marko Terviö

Year of approval 2022 **Number of pages** 52 **Language** English

Abstract

This thesis replicates the study by Vassalou and Xing (2004) which claims that higher level of financial distress risk is the explanation for higher stock returns in small and undervalued firms, a return anomaly identified in financial literature since 1992. They used a financial distress indicator in the US stock market during 1971-1999 to conclude that the higher returns are compensation for carrying financial distress risk that cannot be diversified away. I replicate their study using the same model, but with a smaller sample from US stock market from 2006-2020 and minor differences in data specifications. I find that small firms indeed have higher stock returns, but undervalued firms seem to have unexpectedly lower stock returns, but this latter effect is less precise. I also find no indication that these effects would happen only among high-risk firms or that high distress risk firms in general would have higher stock returns, which was a central result of Vassalou and Xing. After the publication of their study there has been identified both in the area of financial research and other scientific areas a possibility that several of their results might be a result of data snooping. As the authors use methods that have been later identified to inflate the results in this topic area and they also perform multiple tests without correcting acceptance margins, it is likely that their results are not as robust as they claim. Other possible explanations for the absence of the default effect in our sample is the exclusion of smallest firms also known as microcaps, as they also have been identified to inflate results in return anomaly analysis. Related literature has also given several alternative explanations for anomalies that are inconsistent with the classical risk-based explanation for excess returns, and these might affect both the original results and the results of this study. As a conclusion, either distress risk is not able to explain higher returns among larger firms or the model used in this study is able to quantify this risk only among smaller companies. In results of this study the only firm characteristic that indicates higher returns in almost every test even with a higher acceptance margin is small size.

Keywords Anomaly, default risk, bankruptcy, stock returns

Aalto-yliopisto
Kauppakorkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Maisterintutkinnon tutkielman tiivistelmä

Tekijä Jan (Kristian) Koskinen

Työn nimi Distress anomaly revisited – Distance to default in the US stock market

Tutkinto Kauppatieteiden maisteri

Koulutusohjelma Taloustieteen maisteriohjelma

Työn ohjaaja(t) Marko Terviö

Hyväksymisvuosi 2022 **Sivumäärä** 52 **Kieli** Englanti

Tiivistelmä

Tämä tutkielma replikoi Vassalou ja Xingin (2004) empiirisen tutkimuksen, jonka mukaan korkea konkurssiriski selittää pienten ja aliarvostettujen yritysten poikkeavan korkeat osaketuotot, jotka on tunnistettu jo 1980-luvulta saakka. Käyttämällä konkurssiriskimallia Yhdysvaltojen osakemarkkinoilla vuosien 1971-1999 välillä tutkimuksessa tunnistettiin korkeiden osaketuottojen olevan kompensaatiota konkurssiriskin kantamisesta, jota ei voi vähentää hajauttamisella. Replikoin heidän tutkimuksensa käyttämällä samaa konkurssiriskimallia Yhdysvaltojen osakemarkkinoilla pienemmällä otannalla vuosilta 2006–2020 ja pienillä datan muokkauksilla. Tuloksista selviää, että pienillä yrityksillä on odotetusti korkeammat osaketuotot, mutta aliarvostetuilla yrityksillä on yllättävästi alemmat tuotot, vaikkakin jälkimmäinen havainto on epätarkempi. Tuloksista ei myöskään löydy mitään merkkiä siitä, että kyseiset havainnot tapahtuisivat vain korkean konkurssiriskin omaavissa yrityksissä tai että korkean konkurssiriskin yrityksillä olisi yleisesti korkeammat osaketuotot, mikä oli keskeinen tulos Vassalou ja Xingin (2004) tutkimuksessa.

Replikoitavan tutkimuksen julkaisun jälkeen rahoituksen alan tutkimuksessa sekä laajemmassa akateemisessa kirjallisuudessa on tunnistettu monien tutkimustulosten olevan mahdollisesti datan liiallisesta louhinnasta johtuvia. Koska Vassalou ja Xing (2004) käyttävät metodeja, joiden on tunnistettu ylikorostavan tämän tutkimusaiheen tuloksia ja he suorittavat useita tilastollisia testejä ilman että korjaavat luottamusvälejä sen mukaisesti on todennäköistä, että heidän tuloksensa eivät ole yhtä vakaita kuin he väittävät. Muita mahdollisia syitä miksi konkurssiriskin vaikutusta tuottoihin ei havaita tässä tutkimuksessa voi olla pienimpien yritysten rajautuminen otannan ulkopuolelle, koska niiden on havaittu ylikorostavan tuloksia osaketuottopoikkeamien analyysissa. Aihepiirin tutkimuskirjallisuudessa on myös löydetty näille ylituotoille useita klassiselle riskiperusteisuudelle vaihtoehtoisia selittäviä tekijöitä, joiden voidaan olettaa vaikuttaneen sekä replikoitavaan alkuperäiseen tutkimukseen että myös tähän. Johtopäätelmänä joko konkurssiriski ei ole selittävä tekijä suurempien yritysten korkeissa osaketuotoissa tai tutkimuksessa käytetty riskimalli pystyy mittaamaan tätä riskiä vain pienemmissä yrityksissä. Tämän tutkimuksen tuloksissa ainoa selittävä piirre korkeille osaketuotoille lähes jokaisessa testissä ja tiukemmilla tarkkuuskriteereillä on yrityksen suhteellisen pieni koko.

Avainsanat Poikkeama, maksulaiminlyönnin riski, konkurssi, osaketuotot

Contents

1 Introduction

This thesis is in the area of financial anomaly research which investigates sources of consistently higher or lower returns in stock market by looking at common firm characteristics that are observable and quantifiable. In this thesis I replicate a study that claims to have found such characteristic which could explain why smaller and undervalued firms have had historically higher stock returns - financial distress risk, also termed default risk. Several empirical studies have given puzzling results of risky firms providing historically lower returns opposed to what the theory would suggest and the study that I replicate stands out from this almost unanimous crowd of studies by suggesting that distress risk does reward investors and that it can explain other anomalies as well.

The replicated study by Vassalou and Xing (2004) uses Merton's (1974) distance to default model to identify which firms are in risk of defaulting and by comparing portfolio returns and risk levels they find significantly higher returns in firms that the model considers the riskiest. Their results give both a measuring tool for firm level risk in the stock market but also evidence of a positive relationship between risk and reward. The model used in their study, at the time, was the first one to combine data both from the quarterly financial reports and the forward-looking daily data from the stock market to which Vassalou and Xing (2004) credit their more accurate estimates of financial distress. High stock returns have been generally found among small or undervalued companies but from their data Vassalou and Xing (2004) show that both of these anomalies could be explained with their model as well. Riskiest firms still give the highest returns even when cross-inspected within small or undervalued subsets of portfolios. Their results show that financial distress is a risk which cannot be reduced with diversification, and thus provides a premium for investors that carry this risk, and historically carrying the risk would have been a profitable investment strategy.

Replication of the selected study with a slightly newer dataset is mainly motivated by Vassalou and Xing's (2004) strong yet controversial results compared to many similar empirical studies, but also to test the historical experience of anomalies disappearing right after they get published. The broad use of the distance to default model after the study of Vassalou and Xing (2004) (Garlappi, Shu and Yan (2008), Campbell et al. (2008), George and Hwang (2009), Chava and Purnanandam (2010), Eisdorfer et al. (2018), Gao et al. (2018) and Andreou, Lambertides and Panayides (2021)) has also shown it to be a relevant tool in exploring financial distress even today. The dataset used in this replication does not overlap with the one used by Vassalou and Xing (2004) as it spans from January 2005 to December 2020 so this study can be characterized as a follow-up study with a limitation in this new data of not having the smallest firms included. The methodology used in this thesis is as close to theirs as possible, including some possibly erroneous properties discussed later. There were no initial hypotheses set before this study as there is empirical evidence of both negative and positive default risk premium and if the general experience of anomalies disappearing after publication holds also in this case, then we could also expect a null result.

The area of anomaly research has broadened over time, and it has shown several other signals to have consistent return patterns which investors could use as part of their investment strategies as well. Although financial distress is one of the earliest of such signals, together with firm size and market valuation, there still is not consensus whether these anomalies are spurious, interconnected, or separate phenomena. The three anomalies, size, value and distress, nevertheless are inspected together in studies more often than the other anomalies. Some of these studies have found correlations between the factors but have not drawn strong conclusions. However, the results of Vassalou and Xing (2004) at first sight seem to provide a unifying answer to the causal question of the matter. Financial distress appears to be the source single source of at least size and value anomalies and with high statistical precision.

Relying on results from a single study would give a narrow aspect to the topic, especially when several similar studies are available for comparison. To understand both the prevalence of the distress anomaly in the literature, and whether other studies that investigate it have seen size and value anomalies resulting from financial distress as well, requires inspecting these studies from one aspect at time. In the literature review of this thesis, I review empirical studies similar to Vassalou and Xing (2004) first from the aspect of distress anomaly independently, then together with size and value anomalies, to see if other studies have also found evidence of causality or if the connection is merely correlational or non-existent.

The existence or absence of a firm characteristic that systematically signals higher or lower stock returns has far larger implications to the theory than being just a basis for lucrative investment strategy. To appreciate how financial distress is connected to the core of the classical risk and reward framework and the mystery of systemic risk I also review the development of asset pricing theory and the role of default risk in it before reviewing the empirical studies. The empirical studies have also provided explanations for the anomalies that are connected to financial distress but are not within the theoretical framework. Explanations such as investor mispricing of distressed firms, biased analyst forecasts or incorrect handling of delisted firms in the data are gathered to a separate section in the literature review. Evaluating both the results of Vassalou and Xing (2004) and the replication in this thesis with insights drawn from the literature review raises some issues in the methodology but also offers explanations for the differing results in the replication.

The results of this study show that if one had followed the methodology proposed by Vassalou and Xing (2004) among large and mid-cap segments of the US market and invested on risky firms and betted against the less risky ones then one would not have gained any excess returns with the strategy as the average returns are almost the same along the whole risk spectrum. The difference in monthly returns between the riskiest quintile of firms and the least risky ones was 0.041 percent but the standard error is a large 0.556 percent which means that the return difference varied strongly between positive and negative during the time period.

Although in this study financial distress did not explain excess returns, small firms did produce 1.2 percent higher monthly returns compared to large firms and this positive difference seems persistent as the standard error is much lower at 0.4 percent. This size anomaly is well documented in earlier literature, but Vassalou and Xing (2004) found it to be an outcome of higher distress risk. This connection to distress risk seems to have either disappeared over time or it is observable only among the smallest of firms that were excluded from this study.

Undervalued firms have also produced abnormally high returns but in this study, I find this effect to have reversed and similarly as with size anomaly there was no connection to financial distress. Undervalued firms have a negative difference of 0.7 percent in returns compared to highly valued firms, but the accuracy of this result is somewhere between the two previous ones with a standard error of 0.4 percent, which implies that the possibility of this difference being zero cannot be excluded at the confidence level of 95 percent.

A more detailed cross inspection of portfolio returns and Fama Macbeth regressions confirm the same results that small size or high valuation are the strongest indicators of higher returns and the return differences between risky and less risky firm never differ statistically significantly from zero, even among smaller subgroups of firms and time periods, which seems to be a repeating result of this study.

The results are not entirely comparable to those of Vassalou and Xing (2004) since smallest firms are excluded from this sample. This exclusion however might have benefitted us in avoiding a problem identified by Hou et al. (2020) that the smallest firms known as microcaps tend to inflate the size effect if they are used in combination with equally weighted portfolios which were used both in this study and in Vassalou and Xing (2004). The choice of focusing on equally weighted portfolios in their study, as well as in this one, was due to previous literature also using it as main weighing method. The alternative method of using value weighted portfolios was only briefly used in their study and with it their initial results turned statistically insignificant. Similar effect was observed already by Fama (1998) who showed that many anomalous results turn insignificant with value weighting. The idea of equal weighting is also sometimes hard to implement to real markets as institutional investors cannot diversify equally large sums to small firms without ending up as a major shareholder in them and their stocks could be less liquid as well compared to larger firms. Although the whole process of testing multiple times with roughly the same dataset of US market is dubious, Vassalou and Xing (2004) did present impressively accurate and robust results to argue in favor of positive default risk premium. The same tests in our study however give equally consistent, but statistically insignificant results. Either the anomaly was arbitraged away after the publication of their study or the model performs well only if you have the whole market, with microcaps included, in your data.

This thesis is structured so that sections 2-3 are part of a literature review where I provide background information both from common risk factors in general and what empirical studies say about them, but also of the model and methods of analysis used by Vassalou and Xing (2004) and from section 4 onward I describe the data and both the steps and results of my own analysis.

2 Risk, reward and default risk

Understanding the full implications of the results Vassalou and Xing (2004) presented requires backtracking to the core ideas of asset pricing literature. Risk and reward are often mentioned together but the idea of higher risk producing consistently higher rewards is not self-evident. And measuring risk is another topic which at further examination proves to be quite a dilemma. How to measure the risk if we can only observe one of the outcomes? These two topics intertwine when default risk models are utilized in the stock markets and hence, I review next the development of default risk models and asset pricing literature before reviewing studies utilizing these models in section 3.

2.1 Risk in asset pricing

Although data from the financial markets is plentiful, frequently updating and one of the most trusted mood indicators of the expectations in economy, one of the main components of investing is still hard to observe in it – the level of risk. Controlling this risk sometimes means preparing for unpredictable economy-wide shocks but measuring risks in the short run on a more granular scale instead could prove more profitable.

From an economist's point of view a risk neutral player in the stock market has very few optimal choices for a portfolio if he has perfect information and is a profit maximizer, but real markets with their imperfections make it difficult for investors to adjust the amount of risk they take and to observe whether the payoff was in line with the risk taken at the time. Maintaining optimal risk and return level is simple in theory but adding information asymmetry turns it into a continuous balancing act.

Modern portfolio theory developed by Markowitz (1952) states that systemic risk affects every firm similarly, but each firm has idiosyncratic risks that you can diversify away with a large enough portfolio. Diversification allows investors to easily reach optimal risk level with given information, but if one has more information than others about firm specific risks the opportunity arises to sweep the bad risks out of portfolio and invest more on the good ones, yielding higher returns than the market average. This variation in information is one of the main reasons why the theory and empirics are still tangled up in the basic principles.

Accepting that after sufficient diversification the outcome of a stock portfolio is virtually left to the probabilities seems to be even in the present day a hard pill to swallow in some parts of the academia although the idea is one of the most fundamental ones in financial theory. New possibilities to separate winners from losers emerge steadily in financial literature but these anomalies seem to disappear from the markets as soon as they appear in publications (Schwert (2003), Shanaev and Ghimire (2021)). Although anomalies might be short living phenomena, and the studies pinpointing them might tell more about the data itself than of the theoretical principles of finance, compiling their insights reveals interesting peculiarities in the financial markets similarly as behavioral economics does in the field of economics. This branch of literature also pushes its consumers closer to the image of the theoretical investor as it gives more tools to measure risk more accurately as several studies eagerly introduce new models, both simple and complex ones, which can be used to make more sound decisions regarding risk.

In a wholly deterministic universe where all parties have complete information there is no risk as everything has a measurable cause and effect. But in our reality the outcomes which decisionmakers weigh are usually subject to probabilities. Similarly, the signals of a defaulting firm are rarely clear and deterministic, especially if you want to observe them early on, and thus the outcomes can become probabilistic. Firms can face difficulties both on the demand or supply side and exogenous events can have unpredictable consequences to the markets. A company that teeters on the brink of collapse can be propped up artificially with emergency funding or recoverable companies can be tactically slid into bankruptcy. Subjectivity and human decision are part of the default process.

Default and bankruptcies are clearly devastating for a company and its shareholders but for investors who have diversified her investments well, the effect of bankruptcy might not be that large. With sufficient diversification it might even make sense to carry this risk. Classical way to approach risk is the division to systemic and idiosyncratic risk as in the CAPM (Sharpe (1964) and Lintner (1965)) where you distinguish between the risks by whether they can be reduced with diversification (idiosyncratic) or not (systemic). Studies which investigate distress risk have been trying to identify whether distress risk is systemic and if carrying this undiversifiable risk yields a premium as a compensation. This hypothesis that default risk would be systemic fits well to theory and is also intuitive as changes in the market environment can trigger default waves but on the other hand a large company defaulting can also have ripple effects on other companies, potentially causing them to default as well. Even though the systematic risk and thus higher stock returns would be logical to connect to financial distress a share of studies have found financially distressed firms to have lower returns instead, hence the term ‘distress puzzle’.

From the viewpoint of an investor the risk of default is therefore a challenging thing to measure and for this reason there are different types of default risk models focusing either on a narrower aspect of the problem or using proxies to reduce the dimensionality of the problem. To understand the current status of the company one would need to have an insider's view to the company's operations and to get ahead of the curve one needs to also predict the development of the global markets. Default risk is therefore a difficult phenomenon to model but to understand how the models approach differently this problem financial distress needs to be dissected a bit more.

2.2 Financial distress and the risk-based explanation

Financial distress is the more commonly used term for the trigger of default process but compared to the easily observed binary event of bankruptcy, financial distress can vary in its intensity over time, and it is influenced by several factors. To clarify this broad definition Davydenko (2012) separates economic distress (caused by deterioration of market asset values) from financial distress (caused by liquidity shortages) and claims that either one can push a firm into default but usually years after the distress is observed. Economic distress is more connected to the stock markets in which changes in macroeconomy or the firm's area of business affect the valuation of company's asset, whereas financial distress can be more firm-specific as firms choose with their leverage how much they expose themselves to the debt side of the market. Davydenko (2012) also underlines that firms that eventually default show characteristics of both distress types.

Economic and financial distress can be seen parallel to the systemic and idiosyncratic risks in CAPM, and several studies have tested for the systematic nature of default risk in different ways. Some studies such as Shumway (1996) and Dichev (1998) hypothesize that excess returns in itself would be evidence of the risk being systemic as it is priced in returns, but these studies found results that contradict each other, Dichev (1998) finds negative and Shumway (1996) positive returns. Considering also that following studies have not shown steady results of positive or negative returns, but also null results, then the assumption of the risk type being inferable only from the returns seems too simplistic.

Rather than observing the outcome of the risk and return some studies explore the causes of distress instead. Denis and Denis (1995) investigated the reasons why a large share of levered firms encountered financial distress in the 1990s and the 1980s and concluded that firm level characteristics did not explain the different outcomes as much as the macroeconomic changes, which spoke for cyclical and systemic nature of distress risk. An important factor to consider here is that Denis and Denis (1995) observed the importance of macroeconomic changes only among highly levered firms, but Opler and Tittman (1994) see that the vulnerability to macroeconomic changes is strongly connected to firm's capital structure decisions - higher leverage means higher macroeconomic risk, and thus the systemic risk is not entirely disconnected from firm characteristics. Firm level capital structure choices become an important factor also in the detailed analysis of bankruptcies by Asquith, Gertner and Sharfstein (1994) who argue that firm's ability to maneuver through financial distress with actions such as asset sales and debt restructuring are limited if the firm is highly levered.

Capital structure choices then seem to be tied to financial distress, but the idea of leverage amplifying economic risk is not new. Modigliani and Miller (1958) with their classical theorem showed that equity risk should be increasing in leverage and with their assumptions one should observe higher returns in firms that are highly levered and/or distressed but the empirical literature has instead produced many puzzling results of lower returns in such firms. This puzzle does not seem to fit into the framework without a relaxation of the assumptions in Modigliani and Miller (1958) and one example how it could be done is by adding costs to financial distress. George and Hwang (2009) hypothesize that if some firms know financial distress to be costly to them and they choose lower leverage to avoid these costs, then it would be rational for their operational profitability and stock returns to be impacted harder than firms with lesser distress costs that choose higher leverage. With these new assumptions the expected returns would invert, and the puzzling results would then have a risk-based explanation. Tampering with the framework has been less popular choice to explain the puzzle, but what makes George and Hwang (2009) an exception is that they mainly measured financial distress through leverage, a variable that has a clear definition and is measurable. If instead an empirically motivated default risk model were used in this type of analysis, then the model itself would become an easy target for criticism and thus raises the threshold of modifying basic frameworks.

Part of this continuous evaluation of empirical models has been also connected to the investigation of systemic risk. For example, Das et al. (2006), using the probability of default measures from Moody's, finds distress risk levels to be correlated between different types of firms and systematically dependent on the business cycle. The dependency on the business cycle is also observed by Garlappi, Shu and Yan (2008) and

Vassalou and Xing (2004) who see distress indicators peaking during economic downturns. In addition, Vassalou and Xing (2004) mainly argue default risk to be systematic as it has more explanatory power to stock returns than size and value effects even though they acknowledge the discrepancies in earlier literature (Asquith, Gertner and Sharfstein (1994), Opler and Tittman (1994) and Denis and Denis (1995)). All these results, however, are dependent on the model used to measure financial distress which again leaves room for speculation and replication.

The discussion of what is financial distress, how to measure distress risk and is the risk systemic has taken many directions but overall, the literature suggests that there are several good arguments for labeling the risk systemic. Distress risk levels given by the current models covary and are dependent on the macroeconomic changes but there are some idiosyncratic components as well that are mainly linked to the capital structures of firms. Leverage has recently appeared more often in this topic area perhaps due to its better measurability since the fitting of empirical models into the framework of current portfolio theory has proved to be puzzling. The accuracy of predictive models could probably be increased with methods such as neural nets but if the model is not derived from theory, it is hard to see any resolution to these questions that would be agreed upon by academics.

2.3 Default risk models

Distress risk measures are mainly used in the world of credit management but, similarly as in the asset pricing literature, the models have also transited to the equity market mainly to be used in the allocation of stocks into risk categories. Credit rating firms have had special access to a large and detailed datasets of firms and from this data they have developed proprietary models to create measures of creditworthiness such as the probability of the firm migrating from one credit class to another or the probability of default and bankruptcy. The details of these models or the data are usually not publicly available but simpler models with lesser data requirements, such as Altman (1968) Z-score, Ohlson (1980) O-score and Piotroski (2000) F-score, have been published by academics. All of them are very simple to calculate with few financial multiples gathered from financial statements, but the features of the data, such as sparse update intervals or differing reporting practices between countries, are not optimal if you need new information of risks on weekly or monthly basis to manage your portfolio actively. The backward looking and cross-sectional nature of financial reporting has been the main reason why more frequently updating elements from the stock market have been added to some of the models.

Since liquid markets are expected to price assets effectively then the prices can be assumed to contain informational value of the underlying assets. This has been the advantage in models that, in addition to financial statement information, use data from stock or bond market. This allows the model to output estimates more frequently and those estimates can take into account changes in future expectations before they become materialized in the financial statement. Variables such as the measure of market sentiment distress risk can also be proxied from several single sources, such as prices of corporate bond spreads or credit default swaps, but in this study models that utilize a combination of signals are mainly inspected.

The earliest model of distress risk is the Z-score of Altman (1968) who used multiple discriminant analysis to assign weights to five different financial ratios based on 66 companies out of which half did go bankrupt during a certain period. The model reached an accuracy level of 95% when using the latest financial statement analysis data prior bankruptcy but the accuracy decreases with longer prediction horizons. Ohlson (1980) criticized how the division to bankrupt and non-bankrupt firms is too crude a way to investigate financial distress and he also presented concerns whether using multiple discriminant analysis was the proper choice in case of bankruptcy predictions as there are several assumptions that need to be fulfilled when matching firms in the data and selecting the variables to be used in the resulting model. To overcome these problems Ohlson (1980) used conditional logit analysis and ends up with a model that utilizes seven different financial ratios and outputs the probability of bankruptcy instead of a score without an intuitive scale. Latest of the score type models is the F-score of Piotroski (2000) who based on previous research selected nine different binary signals which are calculable from financial statements to create a score by which one can rank firms to financially strong or weak ones. The F-score is rarely seen in other studies perhaps because ranking and sorting firms evenly is inconvenient with a model that outputs only 9 different labels, but perhaps also because the model was used by Piotroski (2000) to separate stocks only in the subpopulation of high book-to-market ratio firms and not in the stock market in general.

The score-models were popular in the 20th century but in the early 2000s the models were enrichened with more frequent data from the markets. A white paper of a credit analysis tool used in a firm named KMV was released by Crosbie and Bohn (2003) and soon after, the model was utilized in studies such as Vassalou and Xing (2004). The model most likely had been used in some form already as it is built on Merton (1974) who applied the classical option pricing model to corporate default prediction. The core idea of Merton (1974) was that by making some assumptions regarding bankruptcy and stock price movement one can use the famed option pricing formulas of Black and Scholes (1973) to infer the risk of company's asset value dropping below its short-term liabilities or in essence defaulting on its loans. A more formalized description of the model is covered in section 4.1 of this thesis, but the main difference is that the Merton model and its variations use daily data from the stock market in addition to the quarterly updating financial statement data which the score-models relied only on.

As different types of distress risk models were available the comparison of them became a topic of interest. Tinoco and Wilson (2013) refer to two comparative studies of distress risk models to show that ranking the models is not easy. Hillegeist et al. (2004) see Merton model more accurate than O-score or Z-score, whereas Agarwal and Taffler (2008) see these model types almost equally accurate. Which type of model is better might be eventually the wrong question as Hillegeist et al. (2004) also note that the less accurate Z-score, O-score and a few additional market variables can add to the explanatory power of the Merton model. Das et al. (2009) also compare the accuracy of score models to the Merton model and find them on par with each other in terms of accuracy and note that the information from the market and accounting metrics are complementary in this purpose.

Evaluating the accuracy of Merton model only in terms of default prediction has also yielded interesting results. Hillegeist et al. (2004) evaluate that by itself the Merton model is not sufficient for assessing the real probabilities of bankruptcy and this is agreed by Bharath and Shumway (2008) who additionally see that a reduced form of the model can perform equally well. Vassalou and Xing (2004) similarly in their own test of Merton model accuracy see the model to be accurate enough for the purpose of ranking firms rather than assessing the real probabilities. They see the default risk estimates to be more accurate than regular volatility measures in terms of distinguishing delisting firms and that the average default risk level rises sharply few years prior the delisting for defaulting firms and remains flat for firms that do not default. The model appears to be a helpful indicator of distress even with simplifying modifications, but if the model is required to estimate real probabilities, then a more accurate model would be preferrable.

An alternative for the option pricing models are hazard models which also employ variables both from financial statements and the markets. The hazard models assume a logistic distribution for the marginal probability of bankruptcies and are based on logit regression. According to Bhararath and Shumway (2008) when comparing the predictions of logit model and Merton model in the highest quintile of default probability the models appear to be equally accurate. Campbell et al. (2008) who use a hazard model to investigate the distress risk anomaly do not specifically characterize Merton model as an equal to their model, but argue that adding the estimates of Merton model to their hazard model adds only relatively little information to it.

The Merton models have nevertheless been favored in studies of distress risk. Studies which either use precalculated estimated default frequencies (EDF) from KMV or any variation of the original distance to default model include Vassalou and Xing (2004), Garlappi, Shu and Yan (2008), Campbell et al (2008), George and Hwang (2009), Chava and Purnanandam (2010), Eisdorfer et al. (2018), Gao et al. (2018) and Andreou, Lambertides and Panayides (2021). Not all authors however use this model as their main approach, Campbell et al. (2008) and Aretz et al. (2019) for example use the distance to default model merely as a benchmark to the hazard model and Chava and Purnanandam (2010) equally test their hypothesis both with the distance to default model and similar hazard model as Campbell et al. (2008).

The way of utilizing the models also varies slightly between studies. Andreou, Andreou and Lambertides (2021) shows with a logit regression that if you use changes in the level of default probability instead of the absolute levels given by the Merton model, one can quite accurately predict stock price crashes. Sudden changes in risk level appear to be better indicators than the risk level itself. Andreou, Andreou and Lambertides (2021) however do not test whether using the 3-month change transfers to portfolio returns, but they give an example how the output of the model can be utilized in different ways.

The older score-models have not been entirely replaced by market-based models even though integrating the daily market information with sentiment and expectations in it is a significant leap. Default, however, seems like a tough phenomenon to predict and none of the models can provide sufficiently accurate predictions for it as the studies show. The empirical studies of anomalies have nevertheless used these models more than a few times and shown that even if default prediction is a difficult task the models can still find differences in stock returns. The question remains how far the models might be allowed to deviate from the requirements of financial theory for this purpose. Tinoco and Wilson (2013), for example, use a neural network as a comparison for model accuracy and even though deep learning can bring us better accuracy than structural models, developing theory with them is much harder. As default risk has been split into different parts in theory perhaps an ensemble model that would consist of structural models accounting for these different aspects could be the next step. If we follow Davydenko (2012) in separating economic distress from financial distress, then Merton model and hazard models with macroeconomic variables could capture the risk of economic distress whereas models that focus on firm level accounting variables could capture the financial distress.

Next, I reconcile how these default risk models have performed in finding differences in stock returns and how controlling for the distress risk has affected the previous known anomalies size and value effect. The original hypothesis was that distress risk should be at the root of both these anomalies, so we should not observe size or value effect if we control for distress and the risk should be systemic. Stocks ranked high in distress should provide excess returns in the long run.

3 Distress, size and value anomalies

In this section I review both how the anomalies are related to the broader asset pricing literature and what empirical studies have to say about the three anomalies that are inspected in the paper of Vassalou and Xing (2004). The anomalies are a central driver in the development of asset pricing literature as they propose for testing new aspects that the financial theory cannot explain. Default risk is one of the anomalies which has had several attempts to explain the puzzling results that it has produced but the effects are usually hard to single out. For that reason I review same studies similar to Vassalou and Xing (2004) separately from the aspect of these three anomalies; size, value and distress.

3.1 Anomalies in asset pricing literature

Asset pricing is divided into two main categories, debt and equity, both which try to estimate the current price of an asset. On equity side the main task is to find a reliable estimate of the value for firms and if they are found to be under- or overvalued in the market then there is an opportunity for profit by buying or selling the stocks of those firms. This pricing process is not static, since the current value of a firm also includes the expected sum of all future earnings and as expectations do change continuously, so does the price. Debt side of asset pricing is similarly concerned with the future of the firm, but more important than the expected amount of earnings is the risk that at some point earnings would sink below the level of due payments and the firm defaults on its debt. This default risk and the models that try to capture it have transitioned over time from the credit side of asset pricing literature to the equity side.

Evaluation of creditworthiness has always required tools such as financial ratios and rules of thumb, but during the 1960s when researchers started to distinguish which indicators were the most effective signals of corporate failure (for example Beaver (1966)) early versions of default risk models started to appear in academic literature. Early models such as Z-score of Altman (1968) were not immediately adopted to asset pricing theory until the anomalies were identified in the 1980s. In between those times a large part of the research regarding bankruptcy prediction and default risk was motivated primarily by bond pricing (e.g. Yawitz et al. (1985), Rodriquez (1988) and Duffee (1999)), since the payoff of the bond is directly connected to solvency of the company, whereas asset pricing theory instead focused on measuring the riskiness of the stock market mainly from historical stock price movement and some macroeconomic variables (see Sharpe (1964), Lintner (1965) and Ross (1976)).

Examination of several anomalies started to appear in the asset pricing literature during the 1980s when the current framework that was mainly dominated by models such as CAPM failed to explain why some observable characteristics of a company, such as small market capitalization (Banz (1981)) or relatively low market value compared to company's assets (Rosenberg, Reid, Lanstein (1984)) produced consistent and abnormally high stock returns when they were used as a basis of portfolio formation. These anomalies guided the asset pricing theory to a multifactorial approach of identifying which source of risk produces high returns instead of the classical division between systemic and idiosyncratic risk.

Large part of modern asset pricing research still derives from the ideas of Fama and French (1993) who incorporated these anomalies in their multifactor model of portfolio returns and suggested that there is an underlying source of risk that drives all of them – financial distress. Although Fama and French were not the first ones presenting this idea (for example Chan et al. (1985) and Chan and Chen (1991) found a strong connection between size effect and indicators of financial distress such as leverage and bond spreads) their contribution gave both a more concise methodology for measuring sources of portfolio returns and at the same time accounted for the recent anomalies. With these developments the asset pricing literature transited to the theory of common risk factors that persists to this day, but it is still under debate what these common risk factors actually are.

Fama and French (1993) left open the mechanism how financial distress transmits into the observed size and value anomalies, and this was partly due to the opaque nature of firms' true financial status which requires a model to identify and quantify it. This gap created by the missing mechanism produced a surge in empirical studies using differing models of financial distress to test the hypothesis that firms in danger of defaulting yield greater compensation for their stockholders for carrying undiversifiable risk. Meanwhile the broader asset pricing literature continued to diverge to new form of anomalies, such as the momentum which is observed from the trend in recent stock price movement (Jegadeesh and Tittman (1993)), instead of pure firm characteristic.

Although there is the possibility of financial distress being at the root of these anomalies, hardly any of the studies that investigate default risk and the stock market have produced results that would have moved asset pricing literature closer to a more simplistic answer. Instead, they have yielded a larger variety of explanations, such as the behavior of investors or information asymmetry, all which have raised doubts if distress risk is just another spurious effect, such as size and value were claimed to be, or if the proxies for distress used in these studies are merely inaccurate. Even though these empirical anomaly studies can be overall characterized ‘mixed and inconclusive’ as Ali and Bashir (2021) put it in their review, reconciled results show interesting dynamics between the anomalies and presents notable criticism or adjustments to the framework in which we examine risk and return.

3.2 Empirical studies of anomalies

This section focuses on three anomalies that are observed directly from firm characteristics. Size and value anomalies, and how they relate to the anomaly of main interest in this study, the distress risk puzzle. Size and value are one of the earliest anomalies in the literature and one reason why they are referred to more often than anomalies that followed (e.g., momentum or low volatility) could be that size and value are the parameters of the famed three-factor model of Fama and French (1993). Although momentum factor was added as the fourth factor to the model by Carhart (1997) momentum is not connected similarly to firm characteristics and thus harder to connect to the risk-based explanation the investigators of distress anomaly are working with. The momentum factor was also replaced more recently by two firm characteristics that can be directly observed, the return spreads between profitable and unprofitable companies (RMW) and conservative and aggressively investing companies (CMA) in the five-factor model by Fama and French (2015) but as majority of the distress anomaly studies covered in this study had been conducted before its introduction the three factors have determined the scope of most studies.

The next sections focus on studies that are done mainly from the aspect of distress puzzle but also cross-inspect size and/or value anomalies. The reason for this is that there are already meta studies for single anomalies (see for example Van Dijk (2011) for size effect) and several studies usually review the literature only from the aspect of maximum two anomalies based on what their research question is. Composing results from all three aspects (size, value and distress anomaly) seems necessary since many studies show these factors to be correlated but the question of causality remains ambiguous. I review studies that have used one or more default risk model in their analysis and inspected at least two of these anomalies. Most relevant studies of default risk should fit to this categorization as the research of these anomalies is heavily interconnected and the measurement of default risk always requires some sort of model. Early studies that proxy financial distress with tools from the debt market, such as bond spreads and credit ratings, are excluded, but as the current literature has moved to use information from the equity market and macroeconomy, examining the results from studies using debt market proxies can be excluded.

First, we focus on the prevalence of distress anomaly in the data, the direction of the default effect, but also the models used to determine whether the results are dependent of the model. Same studies are then inspected from the aspect of size and value anomalies with the focus being in whether the anomalies persist or disappear when financial distress is included.

3.2.1. Default anomaly and default risk models

The first connection between distress risk and stock returns using a distress risk model was found by Dichev (1998) and Piotroski (2000) whose results showed that riskier firms produced lower returns if the risk was measured by score-models based on accounting variables (F-, O- and Z-score). Although the negative returns were puzzling, following studies credited these results more to the value anomaly rather than financial distress. Piotroski (2000) found the negative returns only among value firms with high book-to-market-ratio and Griffin and Lemmon (2002) argued that the proxy measure of financial distress used by Dichev (1998) is possibly driven by value anomaly. Griffin and Lemmon (2002) themselves did not find consistently higher or lower stock returns in distressed firms and thus favored value anomaly as a better explanation. The models based on accounting variables thus introduced puzzling results that did not fit into the risk-based explanation of distress.

The puzzle soon received countering evidence as Vassalou and Xing (2004) made a bold claim that financial distress explains both size and value anomalies, as suggested by Fama and French (1993), since their study produced strong results of both anomalies appearing only among distressed firms. Garlappi. Shu and Yan (2008) used a similar distance to default model to Vassalou and Xing (2004), which utilizes frequent market data, but did not find supporting evidence of positive returns for distressed firms in general, but lower returns instead. They also extended the framework to consider shareholder renegotiation with which their negative result, and possibly earlier puzzling results as well, could be fitted to the theoretical framework. The intuition is that some companies have a better stance than others to renegotiate with debt holders and to make a recovery after defaulting on their loans. The shareholders in the latter type of firms are in real danger of losing their equity and should be rewarded for higher risk than the shareholders in companies that are likely to recover. Their empirical results were intuitive but as the renegotiation ability creates another layer that is hard to measure on top of financial distress, they do not draw strong conclusions from their results.

After Garlappi, Shu and Yan (2008) had presented the potential missing piece of the framework, newer models, but also concerns of potential biases both in the data and the methods, were presented. Campbell et al. (2008) found similar results of negative distress risk premium with a new hazard model that they claimed to be equally accurate as the distance to default model, but they also found distress risk as independent from value and size factors. The negative relationship between distress risk and returns is indeed a more common empiric result, but the results can depend also on whether we are comparing expected returns or realized returns. Chava and Purnanandam (2010) using the same models and following Campbell et al. (2008) do similarly find negative realized returns among distressed companies in their data, but they argue that the expected returns for that time period were actually positive and that in the long run the realized returns would have also turned positive. They underline the abnormally low returns in 1980s that might have been caused by both the large number of bankruptcies filed in mid-1980s and changes in bankruptcy laws in the late 1970s which caused institutional investors to avoid risky stocks. This observation is supported by George and Hwang (2009) who do not find excess returns for distressed firms prior to 1980 but a strong negative premium thereafter by using a more classical risk measure, leverage, in their model. Their choice of using leverage is linked to their argument that leverage amplifies the firm's systemic risks through cost of capital and thus is a valid proxy for distress risk. In theory capital structure choices should not affect asset risk but by extending the framework with market frictions George and Hwang (2009) can fit the puzzling results into it.

As majority of the studies use data from the US market between 1960-2000, the results are, to some level, affected by these surprisingly negative returns of 1980s and especially in studies, such as Dichev (1998), where majority of the data is from the 1980s the observations made by Chava and Purnanandam (2010) and George and Hwang (2009) should especially be considered when reviewing their results. Studies with post 1980s data tend to have a more international aspect, sometimes motivated by another hypothesis presented by Chava and Purnanandam (2010) that distress anomaly exists only within US markets. Gao et al. (2018) and Eisdorfer et al. (2018) by using the distance to default model find some support for the hypothesis as the anomaly is most persistent in the US, but they would also add western European countries to markets where a negative distress risk premium is present as they find abnormally low returns in these countries as well. Aretz et al. (2019) focuses on these same developed countries (US excluded) and although they confirm the existence of the anomaly in these countries using the reduced from model of Campbell et al. (2008), they find an opposite effect as they observe positive distress risk premium from the same time period. The latest studies mainly agree that earlier literature has formed a consensus that distress anomaly is most likely driven by mispricing. Gao et al. (2018) for example also draws similar conclusion from their result and Andreou, Lambertides and Panayides (2021) who is first to use a direct measure of mispricing also finds the negative distress risk premium only among mispriced stocks.

To conclude, the studies that examine the default anomaly have presented plenty of evidence of the distress premium being negative and the results do not seem to be dependent on the model. Only lately the studies have examined data outside the US market in which the tumultuous period of 1980s might have played role in exaggerating the anomaly in the data. But even in other countries the negative premium is observable although many studies conclude it to be most likely due to mispricing. A surprising result is that the anomaly, although a puzzling one, does not seem to have disappeared after its publication. In the next section we cover shortly the history of size effect and review the results with focus on what has been learned of size effect aside the distress puzzle.

3.2.2 Size anomaly and default risk

Size anomaly was identified first by Banz (1981) and the strength of this anomaly in explaining cross-sectional returns was so powerful that studies such as Lakonishok and Shapiro (1985) concluded that size overrules traditional measures of risk in portfolios, including default risk. Many studies afterwards noted that size anomaly seemed to disappear from US data after 1980s right after its publication (see Schwert (2003), Dichev (1998) and Fama and French (1992)) or observed that it has reversed in some countries (Dimson and Marsh (1999)). Size effect might also be connected to systemic risk as Chan et al. (1985) found macroeconomic variables to have a better explanatory power of the high returns in small firms and Kim and Burnie (2002) found the size effect manifesting primarily during good economic conditions and reversing during recessions. Asthakov et al. (2017) also report in their meta-analysis of size effect a significant publication bias, but both the bias and the magnitude of the size effect has decreased over time. Their best estimate of the annual return difference between smallest and largest quintiles in terms of company size is a positive 1.72%. Although size effect was labeled dead, the topic did not disappear from academic research but remains to this day in a smaller role as a complementary factor to other risk measures.

Dichev (1998) was one of the first researchers to compare size and default risk, but before the disappearance of the anomaly from the data his measure of default risk was not a potential explainer of size effect. Soon after Piotroski (2000) observed that the excess returns from the strategy of focusing on strong firms with lower distress risk concentrated on smaller firms, but the results of Piotroski (2000) are hard to generalize as he studies only undervalued firms. His original argument that one can find higher returns by separating stocks even further from just value anomaly-based strategy could have, in the absence of the default risk model, been argued as well to be achieved by using firm size as the criterion. Not soon after Griffin and Lemmon (2002) used O-score as a measure of distress and although they focused more on value than size effect in their study, from their results can be seen that in every risk category small firms generated higher returns. But even if high risk firms tend to be the small ones there was no clear difference in returns between small firms whether they are high or low risk. Campbell et al. (2008) using a hazard model see small firms generally riskier as they have higher chances of defaulting, a view also agreed by Chava and Purnanandam (2010) and George and Hwang (2009), and this connection becomes more robust on longer horizons and especially after the 1980s when small risky firms have a pattern of low returns and high volatility.

Some studies, such as Chava and Purnanandam (2010) and George and Hwang (2009) control for size and value effects in their regressions but refrain from commenting interaction of size and value with default risk. Few studies comment them briefly but without much discussion. Aretz et al. (2019) for example see that including size and value effect does dilute the positive distress premium by a third but the premium remains statistically significant whereas Andreou, Lambertides and Panayides (2021) only comments the size effect to have the expected negative sign and that it is statistically significant when included as a control variable. Gao et al. (2018) using a global dataset see the negative distress risk premium to be the strongest among small firms, but they do not report separately the estimates for size and value factors in their regressions so evaluating how the factors interact is not possible. Interestingly, Eisdorfer et al. (2018) do not discuss why in their regression results size effect in the US has reversed and is highly statistically significant whereas in emerging countries, where distress anomaly seems to be absent, the size effect appears not to have inverted.

To conclude, small firm size and risk of default are correlated and studies which report their results both with and without firm size as control variable seem to show that firm size often dilutes the explanation power of distress risk. The notion that size effect would have disappeared after its publication seems like something that might have been too hastily determined if you inspect it through the studies of distress puzzle. Van Dijk (2011) in his meta-analysis of size effect also concludes that although the size effect has clearly weakened after 1980s, we do not have enough data to declare the anomaly missing.

3.2.3 Value anomaly and default risk

The second major anomaly is the value anomaly which is based on the return difference between value and growth stocks. Distinction between value and growth is made based on how much the market expects the firm to grow in future so growth stocks generally have a market valuation multiple times higher than their current amount of assets and value stocks are valued close to their current asset base. Zhang (2005) refers to a popular rationalization by Grinblatt and Titman (2001) why the high returns of value stocks are anomalous. Since future expectations have uncertainty embedded in them then with rational expectations investors that invest in growth stocks should be rewarded with a premium as they are subject to more systematic risk. Economic shocks could downgrade the growth outlooks and drive down the valuations based on growth greatly. Value stocks on the other hand would not have the same downside as their valuation is close to their assets and since the value of assets is considered redeemable then there should not be a risk premium.

Zhang (2005), however, challenges this view and argues that value anomaly can be fitted into the neoclassical framework if one considers that it is more costly to reduce firm capital than expand it, meaning that value firms actually carry more risk especially in bad times when heavy disinvestment is required whereas growth firms adjust their future plans downward instead of their capital. Whether the theoretical excess returns are considered nowadays anomalous or not, empiric results show that market valuation has been a significant indicator of stock performance and similarly as with firm size there is a link to default risk. The value anomaly has early on been hypothesized to be a result of systematic bias in the valuations. Lakonishok, Shleifer and Vishny (1994) suggest that value stocks tend to have pessimistic expectations and hence are undervalued and although this mispricing argument was presented specifically to value stocks, it concerns both size and distress anomalies as well.

The discovery of the value anomaly is harder to pinpoint as several metrics can be used to separate growth and value stocks. Basu (1977) and Basu (1983) observed the anomaly using earnings-to-price ratio whereas Rosenberg et al. (1985) used book-to-market ratio, which later became the most used indicator of value stocks. Similarly, as with size anomaly, the work of Fama and French (1993) and their three-factor model spurred further research on the value effect especially because they underlined value to be more important in explaining stock returns compared to firm size. As with size effect the value anomaly seemed to initially disappear from the data after publication (Schwert (2003)).

Out of the studies that used default risk models, Dichev (1998), was again the first to challenge the connection between distress and value anomaly. Although risky firms generally could be categorized as value stocks in his data, the riskiest tail of the spectrum that you would expect to be the most undervalued firms were instead stocks with high valuations. The high returns in value stocks could not be explained with distress risk from the data before or after the disappearance of the anomaly. Piotroski (2000) approached next the value anomaly by hypothesizing that the exploitation of value anomaly could be driven even further if you shed distressed firms from the portfolio. Their approach was more motivated by information gained from a fundamental analysis to separate weak and strong firms better than the market rather than by the risk-based explanation. As they expected, the returns indeed were higher for less distressed firms, which might be seen as a contradictory evidence of distress risk premium, but since they only investigated the returns differences within value stocks and excluded growth stocks, it is hard to deduct anything of the overall connection between value anomaly and distress risk.

Griffin and Lemmon (2002) flipped the design of Piotroski (2000) and inspected the value effect within portfolios of different risk levels and they find the strength of value (and size) effect to be increasing with the level of distress risk. Although Griffin and Lemmon (2002) were first to document a connection between distress risk and value effect, Vassalou and Xing (2004) soon presented a more confident hypothesis that distress risk alone could explain both size and value effects, since with their model and data these effects only appeared among the riskiest firms. Garlappi, Shu and Yan (2008) instead saw some dynamic between value and distress in both ends of the spectrum as they saw that among growth stocks higher risk meant lower returns but in value stocks this distress effect reversed. Noteworthy is that although Garlappi, Shu and Yan (2008) report the results as they had inspected the risk levels within value or growth stocks (similarly as in Piotroski (2000)) they actually inspected value effect within different risk groups similarly to Griffin and Lemmon (2002) and from this aspect their results were similar. The effect this type of two-variable sorting has on empirical results is discussed more in section 3.3.2.

As the connection between distress and valuation was now established, also in empirical studies, Campbell et al. (2008) in their investigation of company failures also saw market valuation and bankruptcy being interconnected, as value stocks in general were more likely to delist due to bankruptcy, but they didn't find evidence that this risk would transmit to higher stock returns and conclude that value effect can't be the outcome of a financial distress. George and Hwang (2009) saw negative returns for high distress risk and although their main point was that capital structure (leverage) can potentially explain financial distress better than default risk models, they didn't see it explaining the value effect, and concluded that these are two different types of risk that leverage and book-to-market ratio measure. To further separate the problem into three components Garlappi and Yan (2011) show that leverage (risk) does amplify value effect, but the relationship of value effect and probability of default (distress) is hump-shaped, so that the value effect diminishes among the most distressed firms.

To conclude, similarly to the size effect, the value effect seems to become inconsistent when you add financial distress into the equation and perhaps for this reason some studies such as Chava and Purnanandam (2010), Eisdorfer et al. (2018) and Andreou, Lambertides and Panayides (2021) merely use it as a control variable but refrain from commenting how default risk and the value effect interact. The absence examining the value effect is a bit surprising as from the results of these same studies one can see a positive (Eisdorfer et al. (2018) and Andreou, Lambertides and Panayides (2021)) or negative (Chava and Purnanadam (2010)) value effect and especially the international study of Eisdorfer et al. (2018) they show that the value anomaly is not exclusive to US markets, but a global phenomenon.

Another layer in this branch of literature are all the other possible factors outside the theoretical risk and reward framework, such as market inefficiencies or irrational behavior, that are presented as potential creators of these anomalies. Single studies generally link these potential explanations to a single anomaly, but when inspected as a group the factors that reappear in other studies can be seen to affect not just one but many of these anomalies. As authors such as Hou et al (2020) also present critique of the general methods used in this form of analysis, that might have led to finding these anomalies it is more sensible to inspect them as alternative explanations for anomalies.

3.3 Alternative explanations for anomalies

In this section, I review how studies investigating distress anomaly have identified and accounted for factors outside the theoretical framework of efficient markets, such as market frictions and behavioral aspects, which might be alternative sources of the anomalies. I also inspect whether these factors have influenced the results when they have been controlled for and similarly, I review how studies have identified general methods of analysis used in this research area which might have biased the results more widely.

Since behavioral factors, such as the prospect theory of Kahneman and Tversky (1979), were early on known to affect the functioning of financial markets, it was a common hypothesis that behavioral elements might be significant factors in studies regarding size, value and distress anomalies. In a more recent meta study, Barberis (2021) shows that this indeed was a valid concern since when the effect of prior losses and gains is accounted for, as in prospect theory, they can explain slightly more than half of the total 23 observed anomalies, even though value and size anomalies are not included in them. Even without involving the behavioral aspect, frictions in the market that interfere with efficient trading are generally known to skew the outcomes the theory would predict. Many studies have later identified that these inefficiencies have been in some cases left unaccounted for, but they also have found flaws in the more specific methods of analysis used in those prior studies in this area. A more general concern of publication bias in the academia burdens also this area of literature since at least the results of size anomaly has been identified to have suffered from this bias (Asthakov et al. (2017)). The literature regarding anomalies balances between trying to find explanations within the framework but actively investigates the flaws in it as well.

We inspect these factors in two categories. First, we review results from behavioral and informational factors such as analyst bias and mispricing and in the second section we review more technical factors such as the sorting method of portfolio analysis, data adjustments and statistical noise in the data which can affect the results.

3.3.1 Market inefficiencies, mispricing and behavioral factors

The most common test among studies of distress or value anomaly is to test for mispricing, a systematic under- or overvaluation of certain types of stocks, but the results are mixed even though several studies have found some sort of connection to mispricing with different types of proxies.

Investor mispricing

Lakonishok, Shleifer, and Vishny (1994) and La Porta (1996) presented critique early on to the argument of Fama and French (1992) that size and value can be fitted in to the rational asset pricing model by showing that either typical mistakes of naïve investors (Lakonishok, Shleifer and Vishny (1994)) or too extreme expectations of analysts (La Porta (1996)) can explain the value anomaly. Together, the authors also specified that at the times of earnings announcements the value stocks repeatedly outperform the market expectations (Porta, Lakonishok,Shleifer and Vishny (1997)). Following these arguments Dichev (1998) in his study of distress risk tested if distressed firms experience a correction to returns after new information of bankruptcy risk is available. He observes a slow adjustment over a long period which implies mispricing to be more credible reason for low returns than systematic risk.

Griffin and Lemmon (2002) also use earnings announcements and they strongly claim that value anomaly is the result of mispricing caused by information asymmetry from low analyst coverage. They also find mispricing to be strongest in the group of most distressed firms during earnings announcement. Griffin and Lemmon (2002) hypothesize that investors expect these distressed firms to catch-up with comparable firms with healthier financial status but are systematically disappointed. Gao et al (2018) similarly find support for mispricing as investors seem to underreact when firms receive negative news. Compared to a group of firms that receive no news at all, this underperformance shows that there is temporary overpricing involved. Gao et al (2018) also see negative distress risk premium to be connected to country level measures of overconfidence and to times of bull market. Campbell et al. (2008), on the contrary, find opposite results to this overoptimism of distressed firms as they see distressed firms to outperform others around the time of earnings announcement. Their hypothesis is that announcing earnings at all is a positive sign for a firm that is near bankruptcy. Andreou, Lambertides and Panayides (2021), who find negative returns in distressed firms, actually pinpoints the earnings announcements of distressed firms as the possible main culprit of investor disappointment to distressed firms as he finds evidence that distressed firms adjust their earnings to display more optimistic view than the reality of their business. Distress anomaly also appears only in the most mispriced selection of firms and after controlling for it with a more direct measure of mispricing than in previous studies they see the anomaly disappearing. Investigating the mispricing argument has thus become very common in latest studies.

Analyst bias and illiquidity

Mispricing by investors can also be a derivative of bias in the forecasts of analysts as La Porta (1996) suggested but fewer studies inspect mispricing from this aspect. Chava and Purnanadam (2010), whose results are reliant on the fact that analysts are not systematically biased to forecast distressed firms' performance upwards, test for this bias by using both the most pessimistic and most optimistic forecasts. The idea is to see if investors automatically discount the forecasts downwards, but they do not see this explaining their results of a positive risk premium.

The absence of analyst forecasts can also complicate the valuation of a firm especially if the firm is in distress. Griffin and Lemmon (2002) do see a correlation with low analyst coverage and their measure of high financial distress, but although they are generally in favor of the mispricing argument, they don't see low analyst coverage being significant enough to explain it completely. The results of George and Hwang (2009) give suggestive evidence to the same direction as they test for mispricing only through analyst coverage, but they do not see it explaining their results as well. Additional factors to consider could be dispersion of analyst forecasts or illiquidity of stocks, both which Garlappi, Shu and Yan (2008) use as proxies for mispricing. Although dispersion in forecasts is an indicator of information asymmetry and illiquidity of stocks is a market friction rather than systematic biases, Garlappi, Shu and Yan (2008) do not see either one of them explaining their results. The illiquidity argument, however, has received supporting evidence later, as Gao et al. (2018) see distress anomaly particularly strong among stocks that are very liquid.

3.3.2 Analysis methods, measurement errors and data adjustments

Through development in research methods, a few methodological problems that can distort the results and potentially inflate them as well have been retrospectively identified. One of the core methods when comparing returns of different portfolios is to sort firms based on one or two variables of interest, for example market value (size) and probability of default (distress), and then compare the returns of sub-portfolios. In his meta-analysis of size effect Van Dijk (2011) refers to the results of Berk (2000) who presents valid concern of this two-way sorting method.

Sorting and statistical noise

The problem, according to Berk (2000), is that when stocks are first grouped by a variable, such as firm size, that is known to have explanatory power in stock returns then the variation in stock returns in each group will be smaller than in the whole sample, but because measurement error is usually independent of these variables and unaffected by sorting, the variation of stock returns gets closer and closer to the level of noise in the data and eventually becomes indistinguishable from it.

This of course leads to more null results when two-way sorting is used, but if the sorting variable is correlated with noise, then we can observe seemingly significant results which are instead made of noise. Lo and Mackinlay (1990) demonstrated this in their seminal work of data-snooping and Arnott et al. (2015) also show that small stocks and value stocks tend to be undervalued because of the negatively skewed noise in their price data and they claim that moderate amount of this noise can indeed create size and value effects. The effect of noise in the results of distress risk papers examined in this thesis is hard to evaluate, but if it is related to mispricing, which studies such as Griffin and Lemmon (2002) has found to be connected to distressed firms, then noise could transmit to the end results as well.

Portfolio weighting

Some methods are known to inflate results instead of diluting them. After sorting stocks by a risk factor the next step is to form portfolios from them with equal weighting or value weighting. Equal weighting in theory seems like the rational choice if one wants to diversify into two different stocks with different market prices. With equal weighting, even though the stocks have different prices, one would invest the same amount of money in them and therefore they would have the same weight in the portfolio. With value weighting the more expensive stock has a larger weight in the portfolio proportional to the difference in prices.

Difference in the weighting method can have substantial effects on portfolio returns and there are couple of known factors that together with value weighting can inflate returns of certain stock types. Hou et al. (2020) show that by excluding from their data firms that are smaller than the 20th percentile of NYSE market roughly two thirds of more than 400 anomalies yield insignificant results. Although they underline the microcaps as the main contributor to skewed results, they also emphasize that both equal weighting and the Fama-Macbeth regression together inflate the problem of microcaps. These problems however have not completely been left unaccounted for in earlier literature. Hou et al. (2020) refer to an earlier observation of Fama (1998) that by using value weighting anomalies seem to disappear. The possibility of data snooping was already addressed in early studies such as Piotroski (2000) and Griffin and Lemmon (2002) and many studies test their main results with both value weighting and equal weighting (Griffin and Lemmon (2002), Garlappi, Shu and Yan (2008), Garlappi and Yan (2011) and Aretz et al. (2019)) or use mainly value weighting (Campbell et al. (2008), Chava and Purnanandam (2010), Gao et al. (2018) and Andreou, Lambertides and Panayides (2021)).

Studies that rely only on equal weighting include Dichev (1998) and Piotroski (2000), who do not comment on the choice, and Vassalou and Xing (2004) who argue their use of equal weighting to report the results to be consistent with previous studies regarding size and value effect, but in a small section of their paper where they compare equal and value weighted results, their results are significantly smaller in magnitude with value weighting. Similar dilution is generally seen in studies that report both weighting methods. As exceptions George and Hwang (2009) perform their regressions on a stock level returns, not portfolio returns, so their results do not require weighting method and Eisdorfer et al. (2019) who use mainly value weighting. As they use international data and they don't want large countries with many companies to dominate the portfolios, they use selectively equal weighting in their inference.

Microcaps

The exclusion of microcaps is not a common practice and although the removal of small stocks might seem like a small factor, the effects might be surprisingly large. Fama and French (2008) report microcaps to account for 60% of individual stocks in the three largest stock exchanges of the US, but the monetary value of these firms is only roughly 3% which leads to an extremely skewed distribution. Campbell et al. (2008) and Chava and Purnanandam (2010) remove stocks valued below 1$, Garlappi, Shu and Yan (2005) remove stocks below 2$, George and Hwang (2009), Garlappi and Yan (2011) and Gao et al. (2018) remove stock below 5$, Aretz et al. (2019) removes stocks below the 5th

percentile and studies that do not remove any stocks based on their price include Dichev (1998), Piotroski (2000), Griffin and Lemmon (2002), Vassalou and Xing (2004) and Andreou, Lambertides and Panayides (2021).

The robustness of the results is hard to evaluate in studies that do not exclude microcaps. Some information can be deduced from the results of Garlappi, Shu and Yan (2008) who report that exclusion of the microcaps affects the equal-weighted results whereas value-weighted results are unaffected. Garlappi and Yan (2011) observe that stocks under the price of 5$ are the ones with highest distress risk levels, which might have implications to earlier empirical results of distress anomaly as well.

Negative book-to-market ratio

Another data adjustment method which is unexplored in the literature but might possibly bias the results as well is the inclusion of firms with negative book-to-market ratio. Dichev (1998) specifically mentions that as his measure of distress (Ohlson (1980) O-score) uses book-to-market ratio as an input, the exclusion of firms with negative ratio might bias the results. However. after testing both with and without said firms in the data the results appear not to be affected by it. Griffin and Lemmon (2002) use the same measure of distress, but they instead exclude negative bm-ratio firms without justification as do Vassalou and Xing (2004), Garlappi, Shu and Yan (2008) and Garlappi and Yan (2011). Campbell et al. (2007) on the other hand adjust downwards extremely high values of book-to-market ratio and replace negative values with small positive values. Rest of the studies (Piotroski (2000), Chava and Purnanandam (2010), Gao et al. (2018), Eisdorfer et al. (2018), Aretz et al. (2019) and Andreou, Lambertides and Panayides (2021)) do not comment the negative book-to-market ratios at all so it is assumable that they are included in the data.

Delisting returns

An important factor as we are inspecting the stock returns of distressed companies is how one should account for delisting returns when companies either are taken to private ownership, or they are delisted for performance related issues such as bankruptcy. In the former case shares can be bought out or swapped into shares of the acquirer and in the latter case the stock might lose its value. Shumway and Warther (1999), for example, show evidence that size effect disappears completely in an often-used NASDAQ dataset when stock returns are corrected to reflect the real losses from delisting firms. Some of the studies that investigate distress risk have accounted for this delisting bias by using the average loss rates of -30 or -55 percent which were estimated by Shumway (1997). These include Griffin and Lemmon (2002), Garlappi, Shu and Yan (2008), and Garlappi and Yan (2011).

Some studies rely directly on the delisting returns that data providers, such as CRSP, include in their stock database or they have accounted for delisting in some other form. Dichev (1997), for example, comments only that delisting returns are included in the monthly returns and Campbell et al. (2008) seem to admit that in some cases the returns of portfolios consisting of distressed firms can have an upward bias as they assume stocks are always sold at the end of the month preceding the delisting. Chava and Purnandam (2010) state that their procedure of handling delisting returns is the same as in Campbell et al. (2008) and Gao et al. (2018) note that they are missing delisting returns in their global data outside of United States so results in these countries are biased upward. Piotroski (2000) on the other hand has a perhaps too optimistic assumption that in case of delisting the return is zero, with no losses. Vassalou and Xing (2004) employ more drastic assumption that in case the delisting results to bankruptcy, then the whole capital is lost (-100 percent return) but they do not comment on if they adjust the returns for other types of delisting's. Studies such as Eisdorfer et al. (2019), Aretz et al. (2019) and Andreou, Lambertides and Panayides (2021) do not specify how delisting returns are handled. The monthly delisting returns can affect portfolio returns and for future tests of anomalies Beaver et al. (2007) suggest doing a robustness check by excluding delisting years from the data.

This research area has experienced continuous improvements which on the other hand is desirable but as almost every study seems to have found a new aspect that was overlooked by its predecessors it is hard to put these studies on the same line and compare their results equally. These studies however haven't been able to show nullifying results for the distress puzzle and it seems unlikely that they will provide a binary answer to its existence in the future. The results can definitely be utilized by those interested in either predicting the markets or participate in them. The takeaways of how to take distress risk into account properly are divided based on whether one is interested in a broad range of firms or a selected few. In general, the consensus seems to be that distress risk cannot explain size and value anomalies or vice versa, but there is clearly a correlation between the factors and the most probable factor that might connect them could be mispricing. Distressed firms seem to have, on average, lower stock returns across different countries, but this might change if you observe distressed firms in subsamples such as within undervalued firms.

4 Replication of Vassalou and Xing (2004)

In this section I walk through both the analysis method used in the study of Vassalou and Xing (2004) and inspect in more detail the default risk model used in it. Then I review the data and how it compares to the data both in Vassalou and Xing (2004) and similar studies and summarize if there are any discrepancies that might affect the expected results of this replication. Then I construct the portfolios, test for differences in average returns and perform regression analysis to isolate the effects of each factor. In the end of this section, I perform a few robustness checks and summarize the results.

4.1 Analysis method and the distance to default model

Vassalou and Xing (2004) have three main asset pricing tests. First, they inspect size, value, and distress anomalies independently by sorting all firms by one of these variables and creating five portfolios based on the ranking. Then the monthly returns are calculated for each portfolio and the difference between the returns of portfolios at both ends of the ranking is tested to see if there is a statistically significant difference and thus an opportunity to utilize a long-short strategy of investing to the more profitable end of the spectrum and betting against firms on the other end.

The second test is to sort each of these portfolios by a second factor and to see if observed patterns in the first test hold if we control for one of the factors. Portfolios that were first sorted by distress risk are then sorted by either value or size and portfolios sorted by size or value are then sorted by distress risk. The comparison of return differences is done similarly as in the first test but now one of the factors is always controlled for.

The third test is to use a regression model created by Nijman, Swinkels and Verbeek (2004) which assigns dummy variables for each unique type of portfolio from the previous tests and by having the portfolios constructed both with dual combinations of the factors and only single factors the method allows to extract the effects of each factor level independently as well as their joint effect. As the data is in panel format and observations are grouped in portfolios, Vassalou and Xing (2004) use Fama-Macbeth regression which usually refers to a two-step process of calculating first the factor loadings of each stock and then the risk premiums for each factor by regressing the portfolio returns for each fixed period. In this case only the second step is employed as we are not using factor loadings but dummy-variables of each portfolio type as independent variables. Fama-Macbeth regression outputs the average estimates of all cross-sectional regressions which are run at each point of time. The standard errors produced by this method are not corrected for autocorrelation or heteroskedasticity which according to Fama and French (1988) is a problem in time series data with longer holding periods. To correct for this the Vassalou and Xing (2004) use common Newey West-estimator to report corrected standard deviations and t-stats.

A central piece in this analysis is constructing the default risk model. The are several variations to the model that originates from the work of Merton (1974) and the option pricing formula of Black and Scholes (1973) and often authors have then used slightly differing terms to distinguish them. The most common ways are to either use precalculated expected default frequency (EDF) values provided by Moody's (Garlappi, Shu and Yan (2008) and Gao et al. (2018)) or use the practitioners version of the Moody's KMV model outlined in Crosbie and Bohn (2003) (Vassalou and Xing (2004), Campbell et al. (2008), Eisdorfer et al (2018) and Aretz et al. (2019)) usually referred as Mertons distance to default (DD/MDD) values. Only known study to use the precalculated values from another empiric study is George and Hwang (2009) who are able to use the EDF values calculated by Vassalou and Xing (2004) as they use the same dataset. Bharath and Shumway (2008), in their evaluative study of distance to default, test both the structural form of the model as in Crosbie and Bohn (2003) but also a simplified version with which one can avoid computationally heavy iterative process included in the structural form by using empirically motivated fixed values instead. As Bharath and Shumway (2008) see the naïve version to be roughly equal in accuracy to the structural model Andreou, Lambertides and Panayides (2021) relies only on the naïve version but Chava and Purnanadam (2010) test their results with both versions. Hillegeist et al. (2004) have also modified the system of equations in the structural model slightly, for example by not using firm level stock returns as expected returns, similarly as Vassalou and Xing (2004) does, but using return on assets instead. Campbell et al (2008) who follow the method of Hillegeist et al. (2004) have also modified the same expected return term to a constant 6% as an empirical proxy instead of a firm specific value so there is small variability in the models even if studies at first glance seem to use the same structural model.

The model used here is the structural model similar as in Vassalou and Xing (2004) but as the authors have not released their code with which they have calculated distance to default measures, but only precalculated values, we closely follow the specifications given by Vassalou and Xing (2004) for the inputs of the model. The calculations presented here follow primarily the notation of Vassalou and Xing (2004).

Distance to default model is an application of an option pricing model and hence comes with several assumptions the target of application must satisfy. The main idea why option pricing process would be equivalent to bankruptcy prediction is that in both cases the goal is to infer the probability that the price of an underlying asset drops below a set level by assuming that this probability is already priced in the current market price and past volatility. In case of option pricing the underlying price is the price of the stock that is traded in the market, and the derivative price is the strike price determined arbitrarily by the participants of option trading. In case of call options, the option itself is valuable if the price of the underlying stock remains above the strike price before or at the expiration. The owner of the buy option can earn by exercising the option to buy the stock at a lower price and selling it immediately at higher market price, but if the stock price had dropped below the strike price, then the option becomes worthless. A situation analogous to this would be if the stock price were the value of firm assets and the strike price is replaced with debt liabilities that are due in time *T*, then a shareholder could be in a similar position as the option trader. If the value of firm's assets drops below the amount of short-term liabilities, then the debtors are entitled to redeem all the remaining assets of the firm and the value of the stock has become worthless. The probability of this happening should therefore determine the value of a share similarly as the value of the buy options is determined by the probability of the underlying stock price dropping below the strike price. A shortened description of the model is the following.

If the asset value of a firm follows geometric Brownian motion:

$$dV_A = \mu V_A dt + \sigma_A V_A dW \quad (1)$$

then the change in the value of firm assets V_A over time t consists of the change given by the drift μ (expected return to assets) and asset volatility σ_A. The direction of this change due to volatility is determined by the standard Wiener process W (random walk).

If the price of a call option is replaced with the value of equity V_E, then the Black and Scholes (1973) call option formula transforms into

(2) $$V_E = V_A N(d_1) - D_t e^{-rT} N(d_2)$$

Where *N* is cumulative density function of the standard normal distribution and

(3) $$d_1 = \frac{ln\left(\frac{V_A}{D_t}\right) + \left(r + \frac{1}{2}\sigma_A^2\right)T}{\sigma_A \sqrt{T}}$$

(4) and $$d_2 = d_1 - \sigma_A \sqrt{T}$$

where D_t denotes the face value of the debt observable from the market, *r* the risk-free rate of debt and V_E the value of equity calculated by multiplying the share price by the number of shares. What we cannot observe and for which we must utilize the iterative process are the market value of all assets V_A and the volatility of these assets σ_A.

The iterative process begins by inputting calculated volatility of the stock price σ_E as a starting value for σ_A, solving for V_A, calculating volatility σ_A from new daily values of V_A and with the new volatility solving again for V_A and so forth until σ_A converges. Vassalou & Xing (2004) after obtaining final estimates for V_A and σ_A calculate the drift μ from mean change of $ln(V_A)$, but in the packaged functions of Christoffersen (2020) the estimation of drift is done in every iteration similarly from change in log returns and the iterations stop only after both σ_A and μ have converged.

After converging we have the final estimates for V_A, σ_A, μ and by adding the level of debt D_t observed directly from the financial statement we can calculate the distance to default estimate from

$$DD_t = \frac{ln\left(\frac{V_A}{D_t}\right) + \left(\mu + \frac{1}{2}\sigma_A^2\right)T}{\sigma_A \sqrt{T}}$$

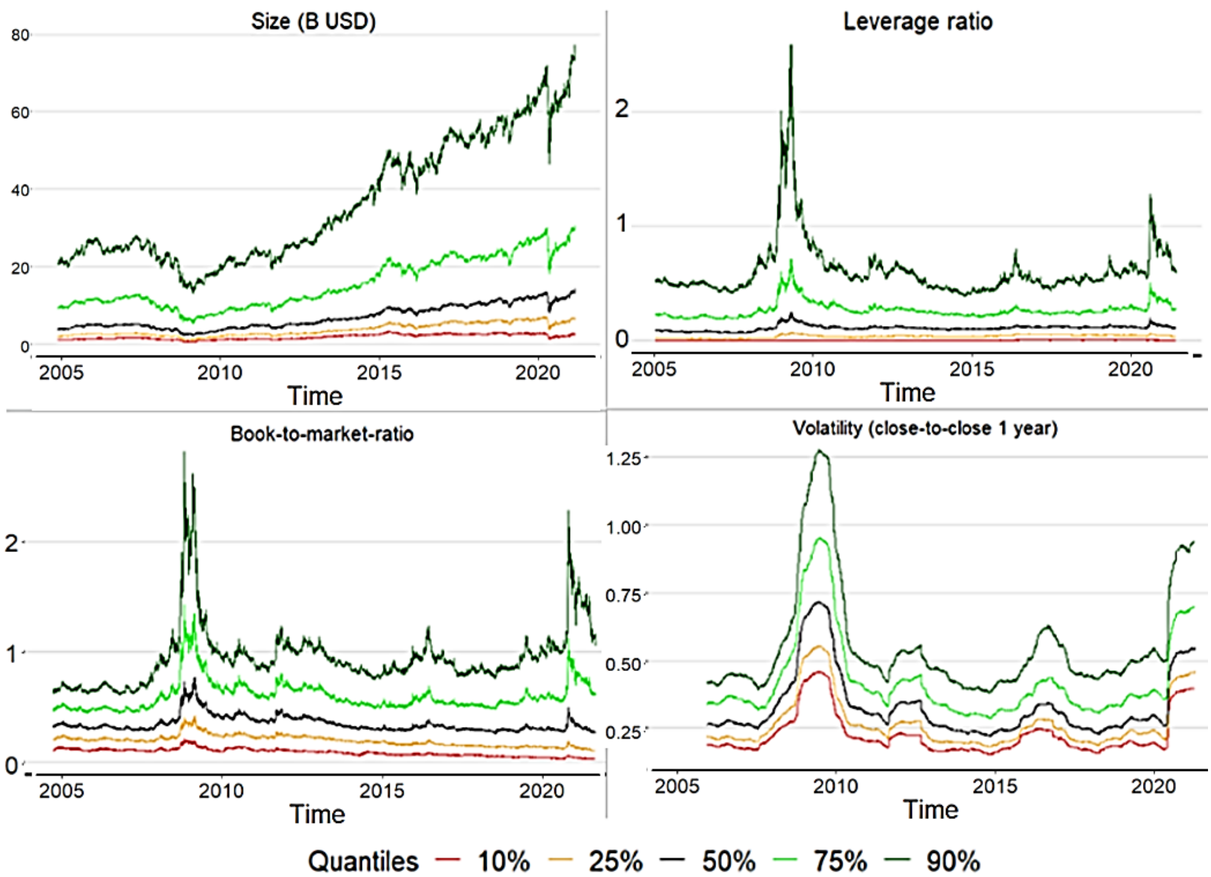
where prediction horizon *T* is one year, level of debt D_t is short-term debt plus half of the long-term debt with daily adjustments, the risk-free rate *r* is the monthly US treasury constant rate and equity volatility σ_E is calculated with common Close-to-Close volatility formula (standard deviation of prices divided by mean price) that is implied to be the method in Vassalou and Xing (2004) although they do not specify it. As a robustness check, volatility was calculated using three different time periods, 12 months prior, 3 months prior and last month. The end results remained similar with all variations.

4.2 Data summary and adjustments

The data in this study is acquired through Standard and Poor's Capital IQ platform and represents Morgan Stanley Capital International (MSCI) USA index from January 2005 to December 2020. At the end of 2020 there were 620 constituents in the index, but as companies might have more than one class of shares and they enter and exit the index, the number of stocks in the index is higher and varies. The stocks that are removed from the index but are still listed in the exchange remain in this dataset. Delisting can happen through a merger, an acquisition or if the company files for chapter 11, a bankruptcy restructuring. The index covers a large 85% share of the market capitalization in US market as most of the large- and mid-cap companies are included. However, this also means that large amount of small companies are excluded. Both the timespan and the time period of this dataset is closer to the international studies of distress anomaly (Eisdorfer et al. (2018), Gao et al. (2018) and Aretz et al. (2019)) who have data from less than twenty years whereas majority of studies that investigate the distress anomaly in US have data roughly from 1965 to 2000. Figure 1 presents descriptive statistics of the data.

Figure 1

Summary statistics of the data with quantiles of market capitalization, leverage ratio, book-to-market-ratio and volatility calculated with 1 year prior price movements.



The data includes daily closing prices of stocks, debt, book-to-market-ratio (BM-ratio), monthly number of shares. I use the monthly US treasury constant rate as the risk-free rate similarly as Vassalou and Xing (2004). The BM-ratio is calculated by comparing the annual book value of equity to daily market value whereas Vassalou and Xing uses semi-annual book values. For debt, similarly as Vassalou and Xing I use annual levels of short-term debt plus half of the long-term debt with daily adjustments.

Vassalou and Xing (2004) have omitted firms with negative BM-ratio, which can occur if the book value of equity is negative due to losses, but the firm is still valued in the market, possibly due to future expectations of returns. Vassalou and Xing do not justify this choice or discuss how large this group is in their sample, but in our sample, there are 207 stocks out of 1205 that receive negative BM-ratios during the whole time period, out of which 120 have negative ratios in at least 12 different months and 87 in at least 24 different months. The model uses market values, not book values, of equity so the calculations should not be affected even if these firms are included, only the ranking of stocks in different portfolios. However, Griffin and Lemmon (2002) also use only nonnegative equity values to calculate the BM-ratio for which reason I also exclude the negative book-to-price-ratio firms in the main analysis. Similarly, as with microcap firms as a robustness check I run the regressions both with and without negative BM-ratio firms.

Vassalou and Xing (2004) also have data of defaults, and in case a firm has defaulted, they assume that investors lose their capital even if some companies still could emerge from bankruptcy. In my analysis no data of defaults is included, but as 391 companies out of 1205 delist before the end of 2020 it is crucial to attempt to identify which of these are due to reasons other than default or bankruptcy. I used public data available on search engines to find out that 25 of these delistings were due to bankruptcies and for almost all other cases the reason for delisting was due to acquisition, merger or renaming of the company. For the firms that delisted due to bankruptcy I similarly assumed that the invested capital was lost, and for those that were privatized the latest closing price was carried to calculate the monthly return. In case of acquisitions which usually are exercised at a price higher than the current stock price, this choice of carrying the latest price should not, on average, inflate the returns calculated in this analysis. There were less than five cases with no information available and for those cases I similarly assumed lost capital as in bankruptcies.

4.3 Single factor portfolios

As a first step I calculate distance to default (DTD) for each firm at the first day of each month and at the same time I assign each firm to a quintile based on three factors: market capitalization (size), BM-ratio (value) and DTD (default risk). The characteristics of portfolios that are formed based on a single factor are listed in Table 1. Vassalou and Xing (2004) only report mean values, not medians, and I follow this choice to make the results comparable. The t-statistics have been calculated from standard errors corrected for heteroskedasticity and autocorrelation by using the Newey-West estimator. According to Lazarus et al. (2018) the standard approach used by Vassalou and Xing (2004) in selecting the lag parameter of the Newey-West estimator has been shown to reduce power of the estimator and thus we follow their recommended general rule and set the lag parameter $1.3T^{½}$. This sets autocorrelation after 18 months to be ignored in our data. I also added leverage-ratio to Table 1 as it has been shown in latest studies to be closely related to financial distress.

Portfolios sorted by distress risk in Panel A have miniscule dispersion in average monthly returns and the average difference between the returns of firms in highest and lowest quintiles is only 0.041 percent with a very high standard error so the effect might go both ways or be zero on average. Vassalou and Xing (2004) observe the opposite as they find at 95 percent level a statistically significant difference of 0.53 percent between high- and low-risk firms when the stocks are equally weighted but with value weighting the difference dilutes to 0.14 percent with a low t-value of 0.46.

Looking at the characteristics of each quintile in Panel A, size and BM-ratio seem to have similar trends as Vassalou and Xing and other authors have observed, smaller firms and firms with high BM-ratio tend to have higher distress risk levels. The BM-ratios also are roughly in the same range as in the sample of Vassalou and Xing, where low-risk firms have a mean BM-ratio of 0.64 and high-risk firms a 1.64. The risk indicator seems to be correlated with the amount of leverage as expected. The difference in the average size of the firms in this sample becomes evident from Panel A as in this sample the average market capitalization ranges from 11.31 to 34.27 billion dollars whereas in the sample of Vassalou and Xing the corresponding values are 2.56 to 5.59 billion. This is most likely a result of this sample excluding the smallest firms.

Next in Panels B and C, firms are sorted by size or BM-ratio to assess their characteristics and whether we can observe the size and value effects. Panel B shows that the size effect seems to be present as there is a 1.227 percent difference between small and large firms with a t-value of 2.961. The point estimate is 0.33 percent higher what Vassalou and Xing observed. Small firms clearly have higher BM- and leverage-ratios and higher level of distress risk measured by our indicator. Although leverage is very high in the smallest quintile, it does not seem to be correlated in other quintiles with size as much as with value or distress risk. We also see that now size varies between quintiles more sharply, which suggests that the risk indicator is not entirely selecting firms based on size.

Table 1
Portfolios sorted by a single risk factor

Average portfolio returns p.m. from January 2006 to December 2020 sorted into quintiles by one of the factors. Standard errors and t-stats corrected for autocorrelation and heteroskedasticity using Newey-West-estimator with lag-parameter of $\sqrt{1.3T}$. Reported values are mean values

Panel A

Portfolios sorted by distress risk indicator to inspect default effect both between quintiles and between the highest and lowest DTD-quintile (High-Low).

			Default risk					
	1 High	**2**	**3**	**4**	**5 Low**	**High-Low**	**Std.error**	**t-stat**
Return	**1.075**	**1.131**	**1.037**	**0.977**	**1.034**	**0.041**	0.556	0.074
DTD	3.176	6.915	9.371	12.011	20.380			
Size	11.312	10.852	14.520	20.877	34.268			
BM-ratio	1.928	0.525	0.429	0.379	0.426			
Leverage	2.202	0.249	0.158	0.104	0.042			

Panel B

Portfolios sorted by size to inspect size effect both between quintiles and between the highest and lowest size-quintile (Small-Large).

			Size					
	1 Small	**2**	**3**	**4**	**5 Large**	**Small-Large**	**Std. Error**	**t-stat**
Return	**1.934**	**1.304**	**0.899**	**0.799**	**0.707**	**1.227**	0.415	2.961
DTD	7.431	10.487	10.690	10.842	12.427			
Size	1.770	4.173	6.943	13.197	61.295			
BM-ratio	8.604	0.720	0.452	0.412	0.393			
Leverage	38.419	0.277	0.266	0.209	0.303			

Panel C

Portfolios sorted by BM-ratio to inspect value effect both between quintiles and between the highest and lowest BM-ratio-quintile (High-Low).

			BM-ratio					
	1 High	**2**	**3**	**4**	**5 Low**	**High-Low**	**Std. Error**	**t-stat**
Return	**0.949**	**0.916**	**1.067**	**1.161**	**1.641**	**-0.692**	0.436	-1.586
DTD	6.283	9.027	11.312	13.233	13.544			
Size	11.499	14.902	18.061	22.287	23.617			
BM-ratio	8.616	0.541	0.349	0.220	0.097			
Leverage	4.146	0.318	0.161	0.108	0.094			

Panel C shows puzzling results. The value effect seems to be negative although with a t-stat of -1.586 the possibility of a null result is fairly large. Earlier literature covered in this study has not produced a negative value effect before. Vassalou and Xing, on the other hand, observed value effect to have a higher estimated return difference than size effect (1.14 percent) and a higher t-stat of 4.588 as well.

Overall, the largest difference in these results is that by comparing the least and most distressed firms in Panel A we did not find any significant difference in returns. This is contrary to the main argument of Vassalou and Xing (2004), that distress risk would be the primary source of risk. Their argument was credible as distress risk consistently seemed to be connected to higher returns on every level in their results, both in aggregate and in smaller sub-portfolios. Regarding distress risk, it is not credible to observe a null result in aggregate and then try to find significant results from subsamples so the expectation for next tests is to try to confirm this null result instead. The only matching result we have, both in terms of magnitude and accuracy, is the size effect. This is somewhat puzzling as well since this sample excludes the smallest firms and Vassalou and Xing (2004) found especially among them high returns of 2.12 percent per month whereas all other size quintiles had average returns varying between 1.16 and 1.28 percent per month. It was expected that all size quintiles in our sample would have had average returns set around that interval but instead they varied from 0.71 to 1.93 percent per month and consistently decreased with firm size. Whether size effect is a significant factor independent of the distress risk depends on whether it is observed in following tests as well.

4.4 Dual factor portfolios

Next, I control for each factor by using two-way sorting and try to separate the effects. First, I control for default effect and inspect both size and value effects in each risk quintile. Then I control for both size and value effect to inspect default effect in size and BM-ratio quintiles. To keep the number of stocks in each portfolio balanced we use sequential sorting similarly as Vassalou and Xing (2004) and first sort the firms based on their distress risk level and then either on size or BM-ratio after which the order is reversed. This sequential sorting also prevents clustering in some of the factors, for example if only the smallest firms tend to have the highest BM-ratios. The results from independent sorting are not reported here as the returns of portfolios were almost identical to these of sequential sorting. Only difference was that small firms and high BM-ratio firms tended to cluster into high-risk quintiles and large and low BM-ratio firms to low-risk quintiles.

Table 2 shows again contradicting results to Vassalou and Xing (2004) who saw a large size effect of 3.82 percent per month with a high t-stat of 9.6 in the high-risk category. In all other risk categories the size effect was insignificant and close to zero. In these results there is statistically significant positive return differences between small and large firms in every risk category decreasing from 1.42 to 0.77 percent per month. Opposed to their results the least accurate estimate in our sample is in the high-risk category.

Table 2

Average portfolio returns p.m. from January 2006 to December 2020 sorted into quintiles by distress risk indicator and firm size. Parentheses distinct portfolios sorted first by the column-variable distress risk and then by row-variable size, which allows to inspect size effect when controlling for default effect. Returns not in parentheses show the default effect when controlling for size effect. Standard errors and t-stats corrected for autocorrelation and heteroskedasticity using Newey-West-estimator with lag-parameter of √1.3T. Correlation matrix of portfolio characteristics calculated with pearson-correlation and reported similarly with parentheses distincting the control variable.

	Default risk								
Size	1 High	2	3	4	5 Low	Whole sample	High-Low	Std. Error	T-stat
1 Small	1.256	1.957	1.951	2.067	1.852	1.934	**-0.596**	0.831	-0.717
	(1.933)	(2.022)	(1.813)	(1.683)	(1.601)				
2	0.921	1.275	1.272	1.351	1.367	1.304	**-0.446**	0.475	-0.939
	(1.185)	(1.235)	(1.052)	(0.944)	(0.953)				
3	0.784	0.714	0.829	0.806	1.003	0.899	**-0.220**	0.382	-0.575
	(0.806)	(0.985)	(0.775)	(0.694)	(0.865)				
4	0.672	0.782	0.862	0.764	0.825	0.799	**-0.153**	0.336	-0.454
	(0.954)	(0.750)	(0.858)	(0.779)	(0.916)				
5 Large	0.348	0.741	0.768	0.792	0.899	0.707	**-0.552**	0.355	-1.554
	(0.512)	(0.662)	(0.692)	(0.774)	(0.828)				
Whole sample	1.075	1.131	1.037	0.977	1.034		0.041	0.556	0.074
Small-Large	**(1.421)**	**(1.360)**	**(1.121)**	**(0.910)**	**(0.773)**	1.227			
Std. Error	0.706	0.228	0.245	0.162	0.223	0.415			
T-stat	2.012	5.970	4.570	5.631	3.464	2.961			

Correlation matrix

	DTD	Size	BM-ratio
DTD	1	(0.278)	(-0.298)
Size	0.347	1	(-0.211)
BM-ratio	-0.415	-0.216	1

When Vassalou and Xing (2004) reverse the sorting order and inspect distress risk in size quintiles they see similarly a high distress premium of 2.23 percent per month with a high t-stat of 5.94. Again, their results are insignificant in all other size quintiles. Our results of distress premium are all insignificant and with negative sign which at least confirms our null result from the first test.

Overall, the default effect seems to be opposite as expected, but the size effect seems to be present in all risk quintiles no matter to which direction we control the effects. Portfolios with smallest firms always seem to outperform the comparison portfolio with all sizes included.

Table 3 shows again that default effect is not significantly different from zero when controlling for value effect and although some of the point estimates have a positive sign they remain quite low in magnitude. The inverse value effect we observed in the whole sample is seen also within each risk category and majority of them have absolute t-stats of over 2 and the rest only around 1.35, but the magnitudes vary between -0.37 and -0.82 percent per month whereas Vassalou and Xing (2004) similarly produced a high 2.55 percent per month positive difference with a t-stat of 9.9 for undervalued stocks in the high-risk category with vastly smaller and less accurate differences in other risk categories almost identical to their size effect. From the correlation matrix we see very large correlations between DTD and size measures, which suggests that to high default risk category has been selection of small firms and large firms into low-risk categories so the differences between high and low risk quintiles might actually be caused by size effect.

Reconciling both the single and dual factor portfolios has given us four main results. First, there is no evidence of general default effect in this sample as sorting firms only by the level of distress risk does not show statistically significant differences in returns. We cannot observe it in smaller subsamples either when controlling for market capitalization and BM-ratio. Second, size effect seems to be observable both in general, as smaller firms have on average 1.23 percent higher returns than large firms, but also within different risk quintiles. Third, we observe also value effect, but to the opposite direction and with lower precision, highly valued firms have higher stock returns than less valued ones. The strength of this effect seems to be slightly dependent on the level of distress risk, but it is hard to determine if this dependency also works the other way around. Fourth, reversing the order by which firms are sorted into portfolios or using independent sorting does not seem to affect portfolio returns by much which was also observed in the data of Vassalou and Xing (2004).

Table 3

Average portfolio returns p.m. from January 2006 to December 2020 sorted into quintiles by distress risk indicator and firm BM-ratio. Parentheses distinct portfolios sorted first by the column-variable distress risk and then by row-variable BM-ratio, which allows to inspect value effect when controlling for default effect. Returns not in parentheses show the default effect when controlling for value effect. Standard errors and t-stats corrected for autocorrelation and heteroskedasticity using Newey-West-estimator with lag-parameter of √1.3T. Correlation matrix of portfolio characteristics calculated with pearson-correlation and reported similarly with parentheses distincting the control variable.

	Default risk								
BM-ratio	1 High	2	3	4	5 Low	Whole sample	High-Low	Std. Error	T-stat
1 High	0.595	1.126	1.004	0.960	0.901	0.949	**-0.307**	0.772	-0.397
	(0.831)	(0.927)	(0.812)	(0.742)	(0.984)				
2	0.950	0.993	0.858	0.864	0.732	0.916	**0.218**	0.374	0.584
	(0.811)	(0.858)	(0.782)	(0.819)	(0.838)				
3	1.207	1.159	0.967	0.923	0.940	1.067	**0.267**	0.294	0.910
	(1.005)	(1.050)	(0.946)	(0.952)	(0.972)				
4	1.244	1.121	1.029	0.952	0.997	1.161	**0.247**	0.322	0.767
	(1.081)	(1.094)	(1.241)	(0.970)	(1.064)				
5 Low	2.162	1.517	1.291	1.276	1.391	1.641	**0.771**	0.477	1.616
	(1.649)	(1.716)	(1.387)	(1.401)	(1.349)				
Whole sample	1.075	1.131	1.037	0.977	1.034		0.041	0.556	0.074
High-Low	**(-0.818)**	**(-0.789)**	**(-0.575)**	**(-0.658)**	**(-0.366)**	-0.692			
Std. Error	0.603	0.285	0.251	0.243	0.270	0.436			
T-stat	-1.356	-2.768	-2.295	-2.714	-1.354	-1.586			

Correlation matrix

	DTD	Size	BM-ratio
DTD	1	(0.905)	(-0.345)
Size	0.897	1	(-0.309)
BM-ratio	-0.428	-0.307	1

4.5 Decomposition of returns

Comparing average returns of different portfolios gives us a good view on how the returns are associated with the three factors. To further examine whether these conclusions persist when we try to further isolate the effects and their interaction, I use portfolio regression analysis method developed by Nijman, Swinkels and Verbeek (2004). Each portfolio type is assigned a dummy variable and then a Fama-Macbeth regression is used to extract the joint and independent effect of each factor level. For the reference portfolios I use the same ones as Vassalou and Xing (2004), low-risk and large size portfolio for the size effect-regression and low BM-ratio and low risk portfolio for the value effect-regression. I also add a third regression where portfolios are sorted by BM-ratio and size to inspect their interaction. The reference portfolio of this regression contains large firms with high BM-ratio. Instead of using tertiles as Vassalou and Xing, I continue to sort firms into quintiles, which brings more explanatory variables to report and thus I only visualize the results here with full regression summary tables in appendix. The choice of using tertiles was done to save space in the study of Vassalou and Xing but according to them using quintiles gave similar results as tertiles. As the data does not include firms with negative BM-ratio, I also report as a robustness check the independent effects with the whole data and with microcap firms excluded.. To evaluate the aspect of systematic risk I also report both independent and interaction effects in different business cycles and through the whole observation period. To save space I do not report the interaction effects for the robustness checks. In total, thirty-five portfolios are formed in each regression, five of each of the two factors included in the regression and twenty-five combination portfolios of those two factors.

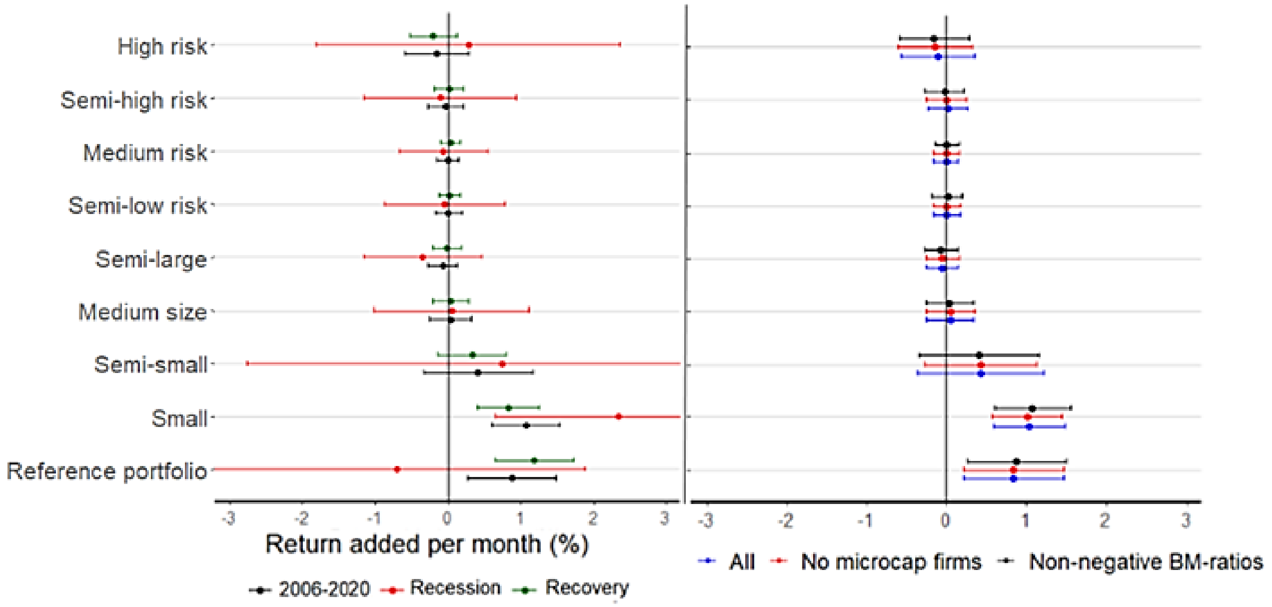
The results are listed in Figures 2-4 and to calculate the return of a combination portfolio I add the point estimates of both factor effects independently and then the joint effect to the return of the reference portfolio similarly to Vassalou and Xing (2004). In this case as the results are quite imprecise, instead of reporting only the point estimates, I report the estimated 95%-confidence intervals and approximate the total returns of the portfolio based on these confidence intervals.

Figure 2

Point estimates and 95%-confidence intervals of how firm size and distress risk affect both independently and jointly monthly portfolio returns. Independent effects compared both between business cycles and between different data filtering options. Joint effects only between business cycles. Returns calculated from January 2006 to December 2020, the reference portfolio contains firms with low distress risk and large size. Microcap firms are firms with less than 300 million USD market cap. Recession periods are between December 2007 and June 2009 and from February 2020 onwards. Recovery periods are all other time periods.

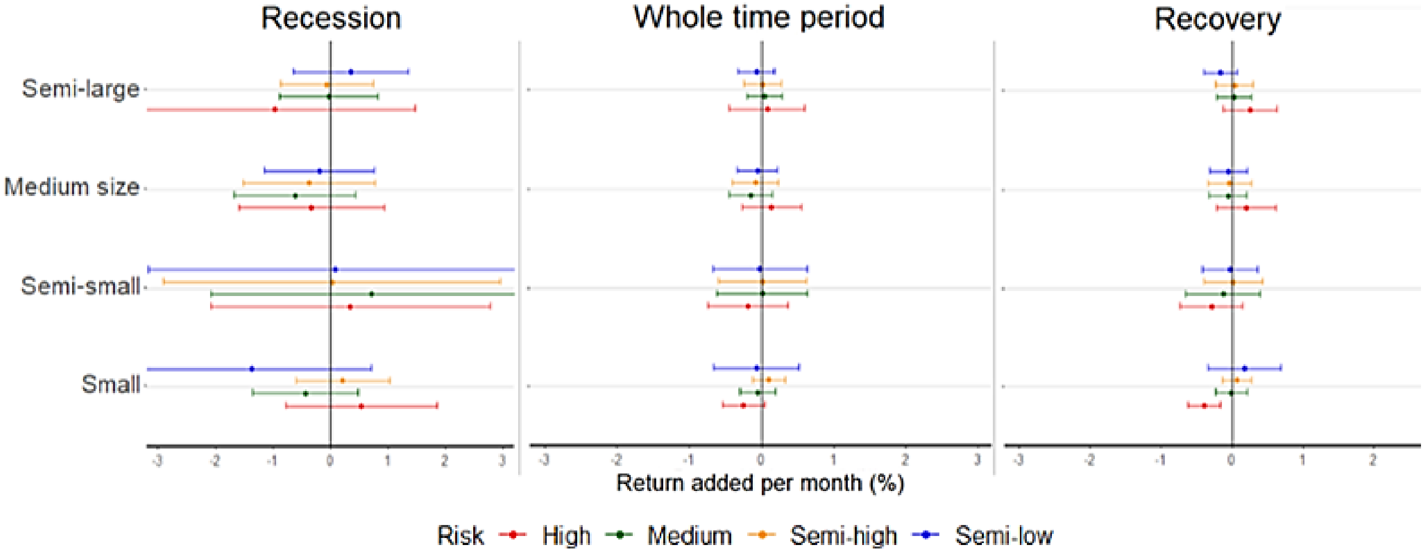
Panel A

Independent effects of size and distress risk



Panel B

Interaction effects of size and distress risk



From the independent effects in Figure 2 we see that generally, there isn't much difference whether we exclude microcap firms or include every firm in the index, but between business cycles there is a large difference in the variance of returns, which is expected. Both the shorter length and the higher volatility of recession periods can explain this. Notable is that the only attribute which adds positive returns in any part of the business cycle is small size and in recession the returns seem to be even larger. Earlier recession periods and size effect were investigated by Kim and Burnie (2002) and their conclusion was that size effect has not entirely disappeared, but it reverses during downturn. This might have been a phenomenon in years preceding our sample or a phenomenon that is dependent on microcaps. Different risk levels do not seem have an effect to returns and although one would expect high risk firms to have produced lower returns during recession periods Eisdorfer et al. (2020) for example argue that there is no underperformance in distressed stocks during downturns if the economic shock affects the entire economy instead of a specific industry. Both the financial crisis and the Covid-19 crisis included in our sample could be considered to have affected the whole US economy which might support their argument. On the other hand, our reference portfolio which consists of large firms with low risk show clear difference between business cycles as they bring positive returns in recovery periods but in downturn the variance increases greatly and both over- and underperformance is possible.

Interaction effects in Panel B show increased variance during recession periods but across the whole observation period all the point estimates are quite close to zero, so these observations are not supportive of the idea of distress risk premium. The only statistically significant interaction is between high risk and small firm size and this combination lowers slightly average returns. Based on these two factors a portfolio consisting of firms in smallest quintile and lowest level of risk would be the one with highest returns in this time period with expected monthly returns between 0.87 and 3.03 percent using the 95% confidence level.

From Figure 3 we see again that in value/risk-portfolios there are no large differences between excluding microcaps or firms with negative BM-ratio. Regarding high risk and high BM-ratio interestingly changing one of them seems to bring highly positive returns in recession periods. The reference portfolio had low BM-ratio and low risk firms so either a combination of low risk and high BM-ratio or high risk and low BM-ratio are the best performing portfolios in a downturn based on this analysis.

Panel B shows that a combination of high BM-ratio and medium to high risk affects negatively to portfolio returns during a downturn with a large magnitude and precision. High risk has positive interaction effects with higher BM-ratio, but as the independent effects of these changes are negative, these effects most likely cancel each other out. The portfolio with highest returns with regards to BM-ratio and distress risk has firms with low BM-ratio but semi-high to high distress risk level with approximate monthly returns of 0.75 to 2.67 percent.

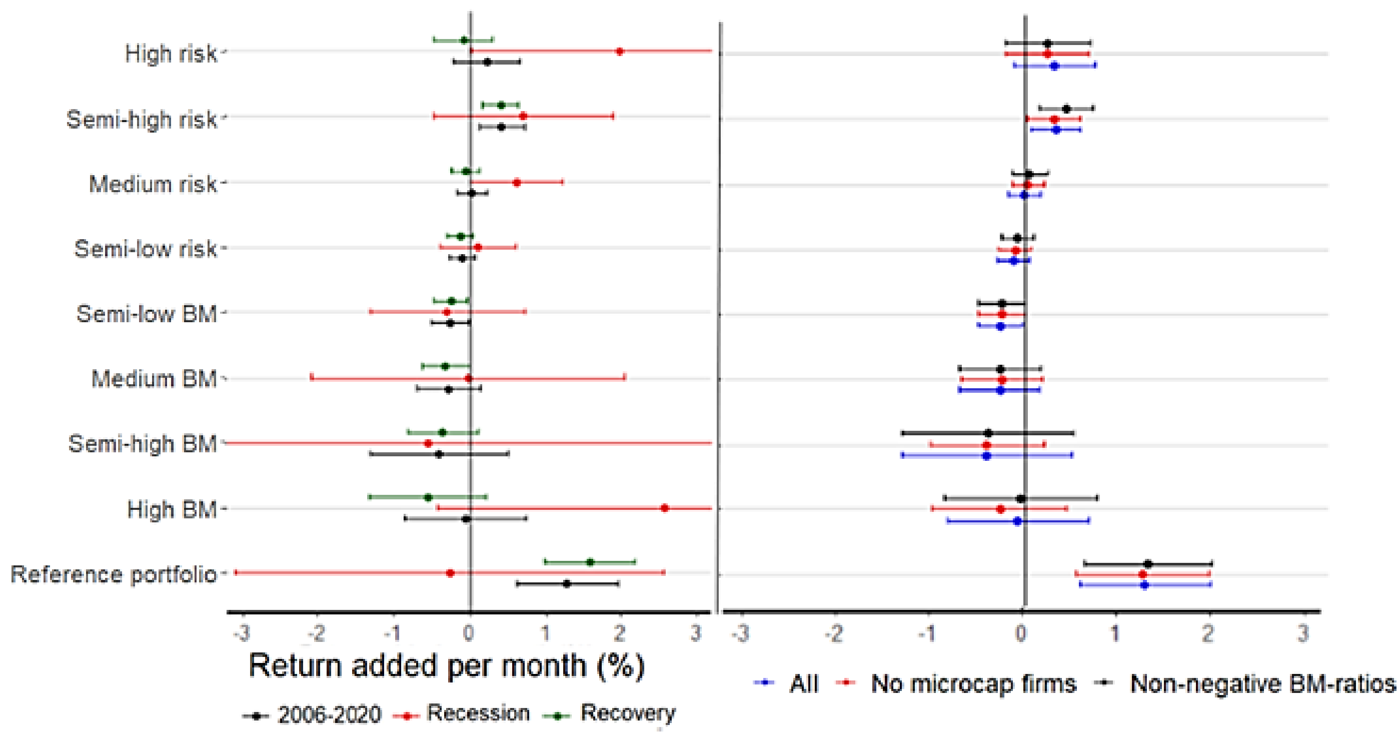
In Figure 4 we have the third regression results from size/value-portfolios which were not included in Vassalou and Xing (2004) but as this study is less centered on distress risk, and gives more room for size and value effects as potential explanations of anomalous returns, inspecting only these two variables seems relevant. From the figure we observe that if I had used the whole dataset without excluding firms with negative BM-ratio these results would have been skewed by the few large positive outliers in our data which in this case would have been part of our reference portfolio. This is seen in the large positive estimate of the reference portfolio and large negative independent effects with wide confidence intervals. This could have been avoided with similar outlier handling as in Campbell et al. (2008) but as I followed the specifications of Vassalou and Xing (2004) we have to inspect the results with this in mind. Otherwise, we see the same pattern observed in Figure 3: small size and low BM-ratio are the characteristics that drive higher returns in portfolios. From Panel B we see that there are hardly negative interaction effects when one considers the whole sample. This suggests that any other portfolio than high BM-ratio and large size is better. The highest returns are in portfolio with small size and low BM-ratio firms with returns between 1.40 and 3.99 monthly returns.

Figure 3

Point estimates and 95%-confidence intervals of how BM-ratio and distress risk affect both independently and jointly monthly portfolio returns. Independent effects compared both between business cycles and between different data filtering options. Joint effects only between business cycles. Returns calculated from January 2006 to December 2020, the reference portfolio contains firms with low distress risk and low BM-ratio. Microcap firms are firms with less than 300 million USD market cap. Recession periods are between December 2007 and June 2009 and from February 2020 onwards. Recovery periods are all other time periods.

Panel A

Independent effects of BM-ratio and distress risk



Panel B

Interaction effects of BM-ratio and distress risk

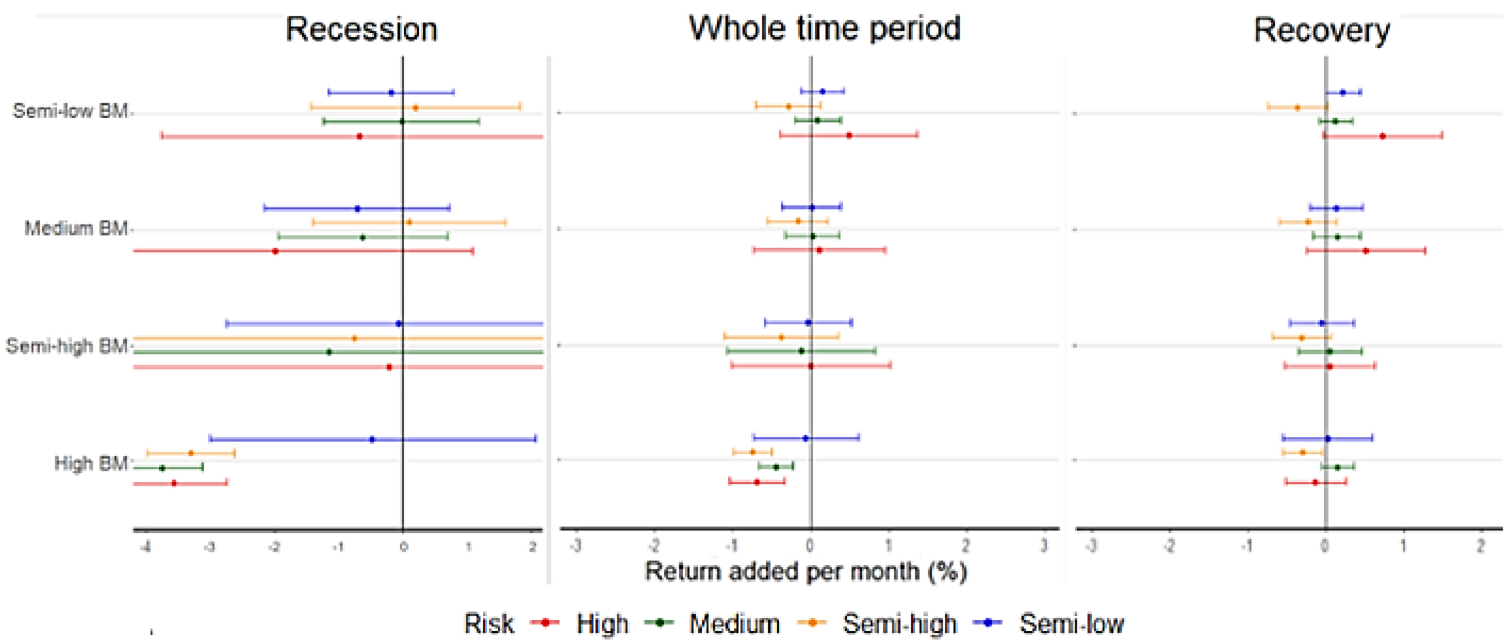
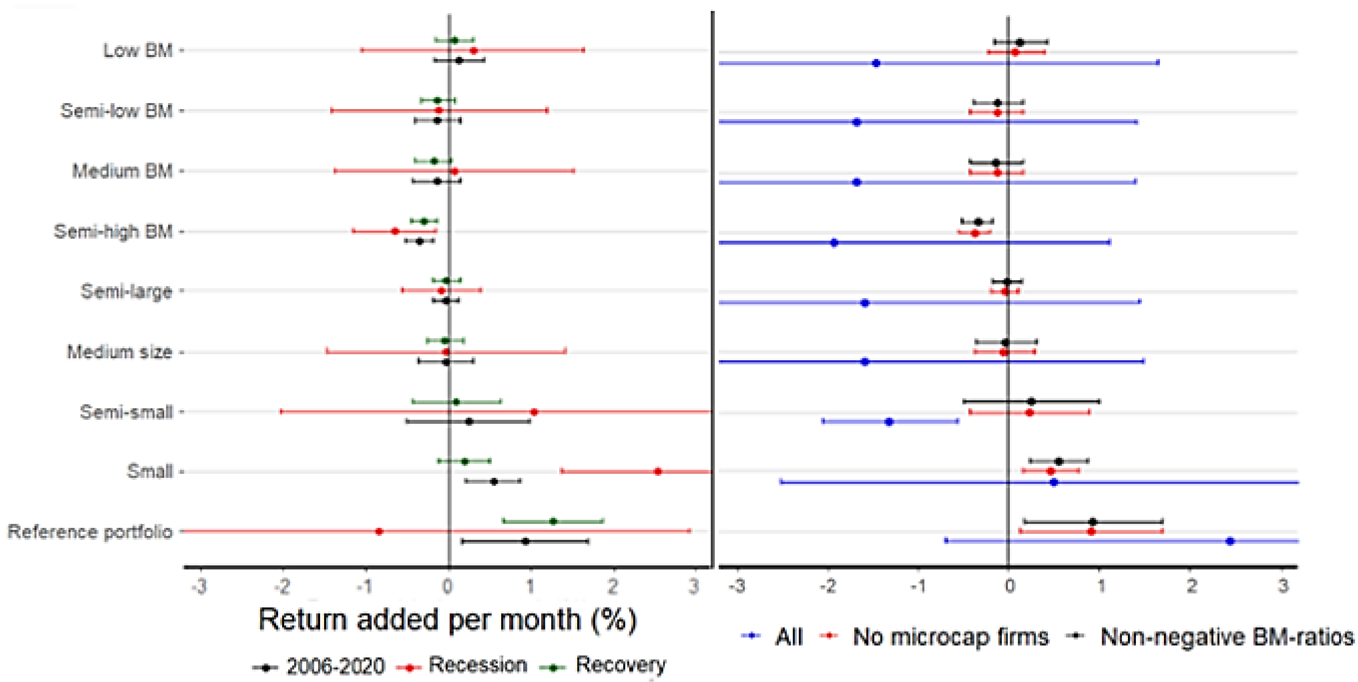


Figure 4

Point estimates and 95%-confidence intervals of how BM-ratio and firm size affect both independently and jointly monthly portfolio returns. Independent effects compared both between business cycles and between different data filtering options. Joint effects only between business cycles. Returns calculated from January 2006 to December 2020, the reference portfolio contains firms with large size and high BM-ratio. Microcap firms are firms with less than 300 million USD market cap. Recession periods are between December 2007 and June 2009 and from February 2020 onwards. Recovery periods are all other time periods.

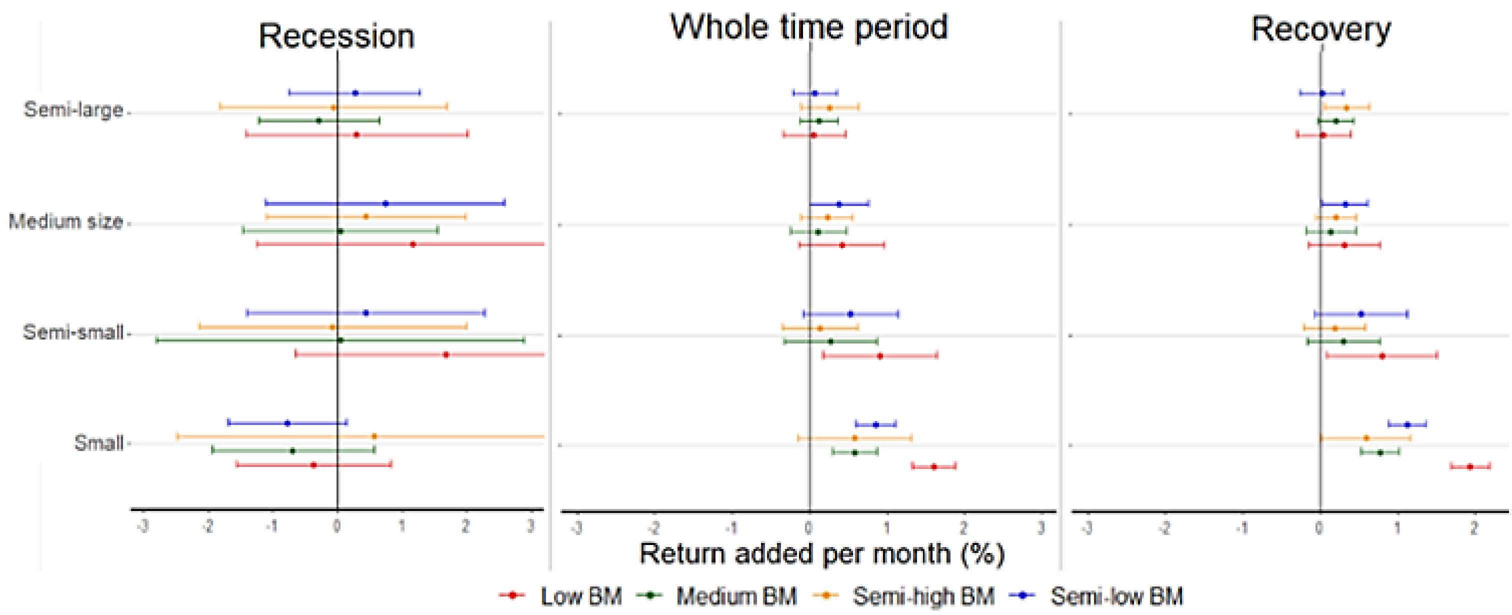
Panel A

Independent effects of BM-ratio and firm size



Panel B

Interaction effects of BM-ratio and firm size



5 Conclusions

This thesis replicated the study of Vassalou and Xing (2004) with data from MSCI USA index from end of 2005 to end of 2020 and opposed to what Vassalou and Xing found, I did not see evidence of size and value effect being the result of financial distress. We did observe that the smallest fifth of firms earned on average 1.2 percent higher returns per month compared to large firms and that low book-to-market-ratio firms had average excess thidifference was found between low-risk and high-risk firms. The differences between portfolios of different distress risk level varied slightly whereas Vassalou and Xing found large, precise and positive 3.7 percent difference in the size effect and 2.1 percent difference in value effect, which suggested that distress risk is the primary cause for both effects.

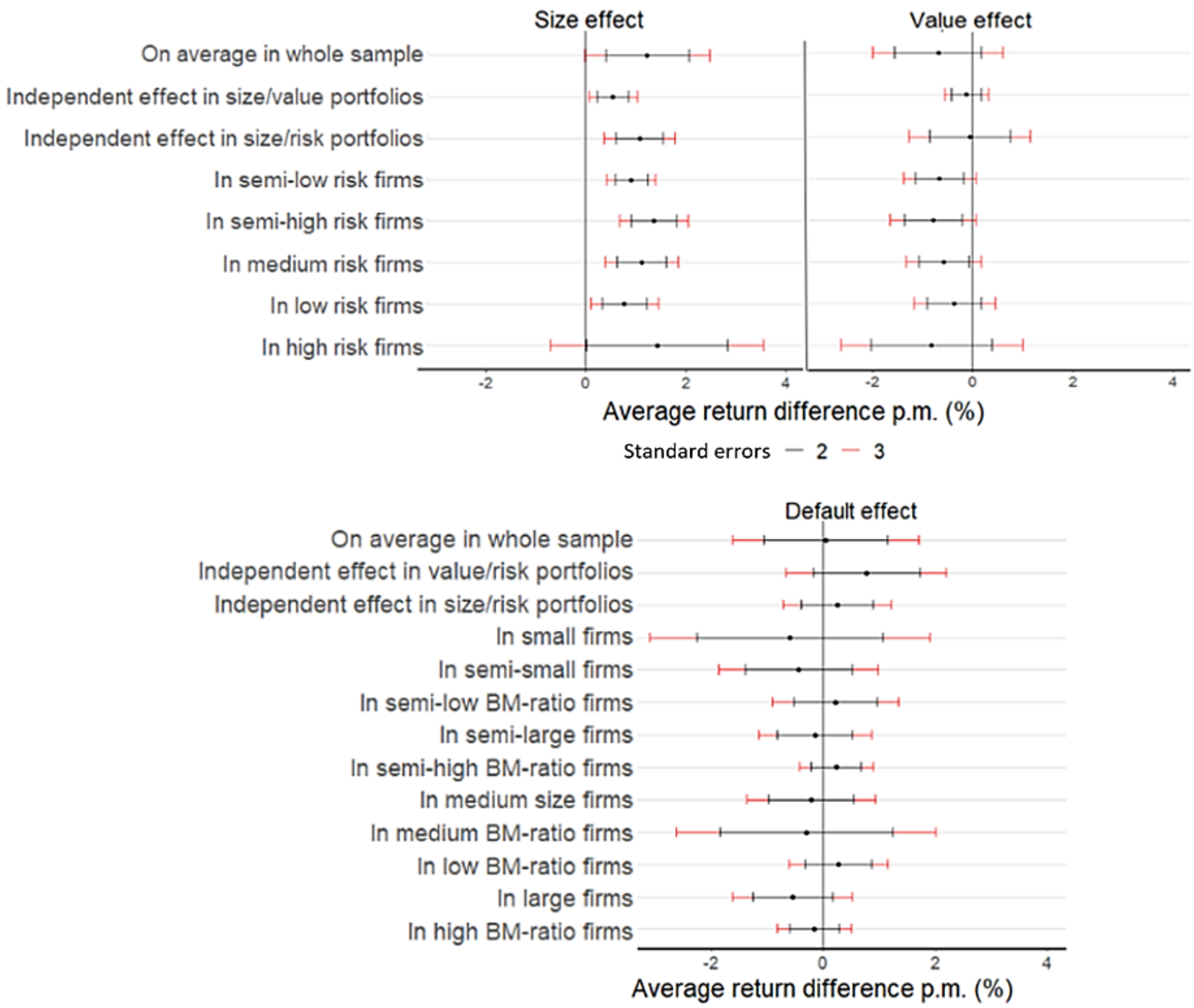
To summarize the multiple tests, Figure 5 presents all our primary estimates of size, value and default effect both with regular 95% confidence interval and a stricter interval of three standard errors. The absence of default effect is evident and a possible explanation for it can be that this sample, which consists of only actively traded large and mid-cap firms with low leverage, has different dynamics than the US stock market overall and default effect can only be observed within smaller companies that are not included in our index. This is supported by the findings of Hou et al. (2020), who find that microcap firms inflate majority of the results in this area of return anomaly research in combination with equally weighted portfolios and use of Fama-Macbeth regressions, all which are present in the study of Vassalou and Xing (2004). Although I did not test whether the default effect would have emerged if microcap firms had been added to the sample, I can conclude that at least the effect is not present in larger firms, or our distress risk indicator is suited only for small and undervalued firms. This would explain why Vassalou and Xing observed the effect only among these firm types.

Another limitation compared to the study of Vassalou and Xing (2004) in addition to the missing microcap firms is the lack of a dataset on defaults. This prevented testing the accuracy of distance to default model in default prediction and although studies such Bharath and Shumway (2008) have concluded that the model and its simplified variations are sufficient estimates for relative distress between companies, these tests are usually performed on datasets with microcaps included in them. Knowing whether distance to default offers sufficient estimates among large and mid-cap segments of firms, as in this data, would have given more insight to whether the model captures the characteristics of small firms instead of actual financial distress. Missing data on defaults also produces another limitation to how delisting firms are handled. I followed the same assumption as Vassalou and Xing (2004) on handling firms that were delisted due to a default but there were five firms to which I could not find information for the reason of delisting. I assumed that they were also delisted due to default and the invested capital was lost, which could slightly bias our results downwards in some portfolios but nevertheless the effect is probably not large. Considering that the literature that followed suggests very different methods to handling delistings, Vassalou and Xings (2004) assumption of losing all invested capital seems rather radical in hindsight and is probably not the right choice for further studies. The missing default data also required an assumption for firms that were delisted for other purposes than default and there is very little f

The only effect that is consistently positive in all our tests is the size effect. The reappearance of size effect is surprising especially in our sample as Knez and Ready (1997) found it to switch direction if you did not include the smallest 1 percent of the firms. Size effect with an average monthly return difference of 1.2 percent also seems surprisingly large if they are compounded to annual percentages since Asthakov et al. (2017) in their meta-analysis estimate annual difference only to 1.7 percent. Studies in this area have often found distressed firms to be overrepresented among small firms and we did observe it in our study as well but the causal link between these two effects seems weak. Either both effects are results of another factor, such as mispricing, or they are independent effects.

Figure 5

Summarization of size, value and default effect estimates from Tables 1-6 with joint effects excluded. Estimated range presented both with 2 standard error and a stricter 3 standard error intervals.



Recommendations and insights based on this study focus thus more on the factors outside the original distress risk premium concept. Both the literature review and our empirical replication show that there are no consistent results, the models used vary significantly and can produce very different estimates. There is also a high number of alternative explanations why earlier studies and our replication could have produced the results they have. Replicating all distress anomaly studies with the proper handling of delisting, mispricing and portfolio weighting would be one step that seems yet to be done but the difficult choice of which default risk model to use in the replication can be a choice that requires taking a stance. If one would prefer as accurate model as possible, deep learning with neural nets would be the obvious choice. This would, however, require abandoning some theoretical frameworks in the process. On the other hand, models based on the theoretical framework might fail to capture the reality of financial distress for as long we don't have unbiased and frequent accounting data from within the firm instead of quarterly reported values. The market proxies have brought the economic distress aspect of the problem available for us and leverage can tell us of the vulnerability to these economic shocks. However, based on the literature review capturing idiosyncratic financial distress is still difficult. If one would adopt the aspect of a reductionist physicist and expect more fundamental laws to be found on a smaller scale, then one should probably approach this question of financial distress by focusing on a more detailed analysis of a firm instead of adding the macro scale of economy and markets into play, as the research has done so far. For an investor practical default risk models are definitely useful, but for research purposes this path might lead nowhere. The literature review revealed also that biases and methodological errors have played a large role in this branch of research and most likely not all of them have been found yet. One thing that has been missing from these studies covered here are the transaction costs. Although in this study all companies are part of an index, and thus most likely traded actively with relatively low trading costs, the costs of smaller firms can be expected to be relatively higher. This would probably decrease the size effect.

The answer to the question whether distress risk is the source of size and value anomalies based on this study and similar ones is a sturdy no and for an investor, who is investing in large and mid-cap firms in US market, distance to default is not a good indicator of excess returns. One should target relatively small firms with low book-to-market-ratio, which in this case provided monthly returns between 1.4 and 4 percent.

References

Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking & Finance*, *32*(8), 1541-1551.

Ali, A., & Bashir, H. A. (2021). Bibliometric study on asset pricing. *Qualitative Research in Financial Markets*.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The journal of finance*, *23*(4), 589-609.

Andreou, C. K., Andreou, P. C., & Lambertides, N. (2021). Financial distress risk and stock price crashes. *Journal of Corporate Finance*, *67*, 101870.

Andreou, C. K., Lambertides, N., & Panayides, P. M. (2021). Distress risk anomaly and misvaluation. *The British Accounting Review*, *53*(5), 100972.

Aretz, K., Florackis, C., & Kostakis, A. (2018). Do stock returns really decrease with default risk? New international evidence. *Management Science*, *64*(8), 3821-3842.

Arnott, R. D., Hsu, J. C., Liu, J., & Markowitz, H. (2015). Can noise create the size and value effects? *Management Science*, *61*(11), 2569-2579.

Asquith, P., Gertner, R., & Scharfstein, D. (1994). Anatomy of financial distress: An examination of junk-bond issuers. *The quarterly journal of economics*, *109*(3), 625-658.

Astakhov, A., Havranek, T., & Novak, J. (2017). Firm size and stock returns: a meta-analysis (No. 14/2017). IES Working Paper.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of financial economics*, *9*(1), 3-18.

Barberis, N., Jin, L. J., & Wang, B. (2021). Prospect theory and stock market anomalies. *The Journal of Finance*, *76*(5), 2639-2687.

Basu, S. (1977). Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis. *The journal of Finance*, *32*(3), 663-682.

Basu, S. (1983). The relationship between earnings' yield, market value and return for NYSE common stocks: Further evidence. *Journal of financial economics*, *12*(1), 129-156.

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of accounting research*, 71-111.

Beaver, W., McNichols, M., & Price, R. (2007). Delisting returns and their effect on accounting-based market anomalies. *Journal of Accounting and Economics*, *43*(2-3), 341-368.

Berk, J. B. (2000). Sorting out sorts. *The Journal of Finance*, *55*(1), 407-427.

Bharath, S. T., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *The Review of Financial Studies*, 21(3), 1339-1369.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy* 81, 637–654.

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In search of distress risk. The *Journal of Finance*, 63(6), 2899-2939.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, 52(1), 57-82.

Chan, K. C., Chen, N. F., & Hsieh, D. A. (1985). An exploratory investigation of the firm size effect. *Journal of Financial Economics*, *14*(3), 451-471.

Chan, K. C., & Chen, N. F. (1991). Structural and return characteristics of small and large firms. *The journal of finance*, *46*(4), 1467-1484.

Chava, S., & Purnanandam, A. (2010). Is default risk negatively related to stock returns?. *The Review of Financial Studies*, *23*(6), 2523-2559.

Chen, N. F., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of business*, 383-403.

Christoffersen B. (2020) Distance to default package https://cran.r-project.org/web/packages/DtD/vignettes/Distance-to-default.pdf

Crosbie, P., & Bohn, J. (2003). Modeling default risk: Modeling methodology. *KMV corporation*.

Das, S. R., Freed, L., Geng, G., & Kapadia, N. (2006). Correlated default risk. *The Journal of Fixed Income*, *16*(2), 7-32.

Das, S. R., Hanouna, P., & Sarin, A. (2009). Accounting-based versus market-based cross-sectional models of CDS spreads. *Journal of Banking & Finance*, *33*(4), 719-730.

Davydenko, S. (2012). When do firms default? A study of the default boundary. In *A Study of the Default Boundary (November 2012). EFA Moscow Meetings Paper, AFA San Francisco Meetings Paper, WFA Keystone Meetings Paper*.

Denis, D. J., & Denis, D. K. (1995). Causes of financial distress following leveraged recapitalizations. *Journal of financial economics*, *37*(2), 129-157.

Dimson, E., & Marsh, P. (1999). Murphy's law and market anomalies. *The Journal of Portfolio Management*, *25*(2), 53-69.

Duffee, G. R. (1999). Estimating the price of default risk. *The Review of financial studies*, *12*(1), 197-226.

Eisdorfer, A., Goyal, A., & Zhdanov, A. (2018). Distress anomaly and shareholder risk: International evidence. *Financial Management*, *47*(3), 553-581.

Eisdorfer, A., & Misirli, E. U. (2020). Distressed stocks in distressed times. *Management Science*, *66*(6), 2452-2473.

Eisdorfer, A., Goyal, A., & Zhdanov, A. (2011). Misvaluation and return anomalies in distress stocks. *Swiss Finance Institute Research Paper*, (12-12).

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of financial economics*, *49*(3), 283-306.

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of political Economy*, 96(2), 246-273.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, 47(2), 427-465.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1), 3-56.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *The journal of finance*, *51*(1), 55-84.

Fama, E. F., & French, K. R. (2008). Dissecting anomalies. *The Journal of Finance*, *63*(4), 1653-1678.

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics,* 116(1), 1-22.

Gao, P., Parsons, C. A., & Shen, J. (2018). Global relation between financial distress and equity returns. *The Review of Financial Studies*, *31*(1), 239-277.

Garlappi, L., Shu, T., & Yan, H. (2008). Default risk, shareholder advantage, and stock returns. *The Review of Financial Studies*, *21*(6), 2743-2778.

George, T. J., & Hwang, C. Y. (2010). A resolution of the distress risk and leverage puzzles in the cross section of stock returns. *Journal of financial economics*, *96*(1), 56-79.

Griffin, J. M., & Lemmon, M. L. (2002). Book-to-market equity, distress risk, and stock returns. *The Journal of Finance*, 57(5), 2317-2336.

Grinblatt, Mark, and Sheridan Titman, 2001, Financial Markets and Corporate Strategy, 2nd Edition (McGraw-Hill).

Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the cross-section of expected returns. *The Review of Financial Studies*, 29(1), 5-68.

Hillegeist, S. A., Keating, E. K., Cram, D. P., & Lundstedt, K. G. (2004). Assessing the probability of bankruptcy. *Review of accounting studies*, *9*(1), 5-34.

Hou, K., Xue, C., & Zhang, L. (2020). Replicating anomalies. *The Review of Financial Studies*, 33(5), 2019-2133.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, *48*(1), 65-91.

Jessen, C., & Lando, D. (2015). Robustness of distance-to-default. *Journal of Banking & Finance*, 50, 493-505.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kim, M. K., & Burnie, D. A. (2002). The firm size effect and the economic cycle. *Journal of Financial Research*, 25(1), 111-124.

Knez, P. J., & Ready, M. J. (1997). On the robustness of size and book-to-market in cross-sectional regressions. *The Journal of Finance*, *52*(4), 1355-1382.

Lakonishok, J., & Shapiro, A. C. (1986). Systematic risk, total risk and size as determinants of stock market returns. *Journal of Banking & Finance*, *10*(1), 115-132.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1994). Contrarian investment, extrapolation, and risk. *The journal of finance*, *49*(5), 1541-1578.

Lazarus, E., Lewis, D. J., Stock, J. H., & Watson, M. W. (2018). HAR inference: Recommendations for practice. *Journal of Business & Economic Statistics*, 36(4), 541-559.

La Porta, R. (1996). Expectations and the cross-section of stock returns. *The Journal of Finance*, *51*(5), 1715-1742.

Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *The journal of finance*, *20*(4), 587-615.

Lo, A. W., & MacKinlay, A. C. (1990). Data-snooping biases in tests of financial asset pricing models. *The Review of Financial Studies*, *3*(3), 431-467.'

Markowitz, H. (1952) Portfolio Selection. *The Journal of Finance*, Vol. 7, No. 1, pp. 77-91. March. 1952.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of finance*, 29(2), 449-470.

Modigliani, F., & Miller, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *The American economic review*, *48*(3), 261-297.

Nijman, T., Swinkels, L., & Verbeek, M. (2004). Do countries or industries explain momentum in Europe? *Journal of Empirical Finance*, *11*(4), 461-481.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of accounting research*, 109-131.

Opler, T. C., & Titman, S. (1994). Financial distress and corporate performance. *The Journal of finance*, *49*(3), 1015-1040.

Piotroski, J. D. (2000). Value investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research*, 1-41.

Porta, R. L. (1996). Expectations and the cross-section of stock returns. *The Journal of Finance*, *51*(5), 1715-1742.

Porta, R. L., Lakonishok, J., Shleifer, A., & Vishny, R. (1997). Good news for value stocks: Further evidence on market efficiency. *the Journal of Finance*, *52*(2), 859-874.

Rodriguez, R. J. (1988). Default risk, yield spreads, and time to maturity. *Journal of Financial and Quantitative Analysis*, *23*(1), 111-117.

Rosenberg, B., Reid, K. , & Lanstein, R. (1998). Persuasive evidence of market inefficiency. *Journal of Portfolio Management* 11 (1984): 9-17.

Ross, S.. 1976, The arbitrage theory of capital asset pricing. *Journal of Economic Theory* 13. 341-360.

Schwert, G. W. (2003). Anomalies and market efficiency. *Handbook of the Economics of Finance*, *1*, 939-974.

Shanaev, S., & Ghimire, B. (2021). Efficient scholars: academic attention and the disappearance of anomalies. *The European Journal of Finance*, *27*(3), 278-304.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The journal of finance*, *19*(3), 425-442.

Shumway, T. G. (1996). The premium for default risk in stock returns (Doctoral dissertation, University of Chicago Graduate School of Business).

Shumway, T. (1997). The delisting bias in CRSP data. *The Journal of Finance*, *52*(1), 327-340.

Shumway, T., & Warther, V. A. (1999). The delisting bias in CRSP's Nasdaq data and its implications for the size effect. *The Journal of Finance*, *54*(6), 2361-2379.

Stephen A Ross, (1976) The arbitrage theory of capital asset pricing, *Journal of Economic Theory*, Volume 13, Issue 3, Pages 341-360

Tinoco, M. H., & Wilson, N. (2013). Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. *International review of financial analysis*, *30*, 394-419.

Van Dijk, M. A. (2011). Is size dead? A review of the size effect in equity returns. *Journal of Banking & Finance*, *35*(12), 3263-3274.

Vassalou, M., & Xing, Y. (2004). Default risk in equity returns. *The journal of finance*, 59(2), 831-868.

Yawitz, J. B., Maloney, K. J., & Ederington, L. H. (1985). Taxes, default risk, and yield spreads. *The Journal of Finance*, *40*(4), 1127-1140.

Zhang, L. (2005). The value premium. *The Journal of Finance*, *60*(1), 67-103.

Appendix

Table 4

Regression summary of portfolios independently sorted by size and distress risk with three different Fama-Macbeth-regressions run in different business cycles. Standard errors and t-stats corrected for autocorrelation and heteroskedasticity using Newey-West-estimator with lag-parameter of 3.

		2006-2020				**Recovery**				**Recession**			
		Estimate	Std. Error	t-stat	Pr(>|t|)	Estimate	Std. Error	t-stat	Pr(>|t|)	Estimate	Std. Error	t-stat	Pr(>|t|)
Independent effects	Reference portfolio (Large & low risk)	0.874	0.306	2.853	0.004	1.185	0.268	4.421	0.000	-0.698	1.292	-0.540	0.589
	Small	1.073	0.234	2.831	0.005	0.824	0.212	3.390	0.001	2.345	0.844	1.160	0.246
	Semi-small	0.411	0.373	2.681	0.007	0.334	0.232	2.549	0.011	0.740	1.748	1.095	0.274
	Medium size	0.037	0.145	0.348	0.728	0.037	0.122	0.384	0.701	0.055	0.532	0.121	0.903
	Semi-large	-0.072	0.104	-0.763	0.446	-0.012	0.095	-0.170	0.865	-0.342	0.406	-0.746	0.456
	Semi-low risk	0.007	0.093	0.099	0.921	0.025	0.068	0.354	0.724	-0.050	0.410	-0.184	0.854
	Medium risk	0.000	0.074	-0.001	0.999	0.034	0.067	0.295	0.768	-0.058	0.301	-0.111	0.912
	Semi-high risk	-0.031	0.120	-0.142	0.887	0.014	0.101	0.082	0.934	-0.107	0.523	-0.101	0.919
	High risk	-0.158	0.220	-0.346	0.729	-0.202	0.162	-0.658	0.510	0.279	1.040	0.116	0.908
Joint effects	Small & semi-low risk	-0.068	0.294	-0.209	0.834	0.176	0.259	0.759	0.448	-1.372	1.036	-0.829	0.407
	Semi-small & semi-low risk	-0.019	0.327	-0.132	0.895	-0.029	0.197	-0.207	0.836	0.082	1.625	0.154	0.878
	Medium size & semi-low risk	-0.058	0.136	-0.435	0.664	-0.049	0.133	-0.393	0.694	-0.201	0.477	-0.386	0.699
	Semi-large & semi-low risk	-0.066	0.125	-0.574	0.566	-0.154	0.115	-1.394	0.163	0.347	0.500	0.824	0.410
	Small & medium risk	-0.046	0.121	-0.138	0.890	-0.007	0.111	-0.025	0.980	-0.436	0.458	-0.274	0.784
	Semi-small & medium risk	0.009	0.313	0.059	0.953	-0.123	0.261	-0.815	0.415	0.699	1.396	1.299	0.194
	Medium size & medium risk	-0.144	0.148	-1.118	0.264	-0.054	0.133	-0.446	0.656	-0.620	0.527	-1.242	0.214
	Semi-large & medium risk	0.041	0.124	0.351	0.726	0.031	0.121	0.285	0.775	-0.032	0.428	-0.067	0.947
	Small & semi-high risk	0.102	0.110	0.315	0.753	0.071	0.102	0.292	0.770	0.210	0.405	0.131	0.895
	Semi-small & semi-high risk	0.016	0.305	0.102	0.918	0.022	0.204	0.135	0.893	0.024	1.463	0.046	0.963
	Medium size & semi-high risk	-0.084	0.160	-0.624	0.533	-0.033	0.153	-0.276	0.783	-0.375	0.572	-0.663	0.507
	Semi-large & semi-high risk	0.016	0.128	0.125	0.900	0.040	0.132	0.364	0.716	-0.070	0.401	-0.121	0.904
	Small & high risk	-0.249	0.144	-0.988	0.323	-0.391	0.116	-1.804	0.071	0.535	0.658	0.484	0.628
	Semi-small & high risk	-0.182	0.276	-0.826	0.409	-0.285	0.221	-1.413	0.158	0.342	1.209	0.378	0.705
	Medium size & high risk	0.144	0.203	0.619	0.536	0.205	0.209	1.053	0.292	-0.339	0.631	-0.328	0.743
	Semi-large & high risk	0.080	0.259	0.340	0.734	0.260	0.187	1.322	0.186	-0.975	1.215	-0.931	0.352
		R^2	0.976			R^2	0.970			R^2	0.980		

Table 5

Regression summary of portfolios independently sorted by BM-ratio and distress risk with three different Fama-Macbeth-regressions run in different business cycles. Standard errors and t-stats corrected for autocorrelation and heteroskedasticity using Newey-West-estimator with lag-parameter of 3.

		2006-2020				Recovery				Recession			
		Estimate	Std. Error	t-stat	Pr(>|t|)	Estimate	Std. Error	t-stat	Pr(>|t|)	Estimate	Std. Error	t-stat	Pr(>|t|)
Independent effects	Reference portfolio (Low BM & low risk)	1.285	0.335	3.833	0.000	1.587	0.295	5.369	0.000	-0.272	1.410	-0.193	0.847
	High BM	-0.060	0.404	-0.136	0.892	-0.555	0.384	-2.764	0.006	2.565	1.501	1.017	0.309
	Semi-high BM	-0.404	0.453	-2.185	0.029	-0.356	0.226	-2.353	0.019	-0.542	2.186	-0.636	0.525
	Medium BM	-0.278	0.213	-2.188	0.029	-0.318	0.158	-2.834	0.005	-0.025	1.034	-0.046	0.963
	Semi-low BM	-0.261	0.126	-3.138	0.002	-0.249	0.108	-2.925	0.003	-0.295	0.505	-1.084	0.278
	Semi-low risk	-0.091	0.083	-1.014	0.310	-0.124	0.087	-1.345	0.179	0.112	0.247	0.394	0.693
	Medium risk	0.034	0.093	0.222	0.824	-0.059	0.092	-0.443	0.658	0.615	0.305	0.959	0.337
	Semi-high risk	0.423	0.146	1.660	0.097	0.408	0.119	2.019	0.043	0.712	0.589	0.596	0.551
	High risk	0.226	0.220	0.412	0.680	-0.086	0.190	-0.206	0.837	1.975	0.979	0.740	0.459
Joint effects	High BM & semi-low risk	-0.059	0.335	-0.175	0.861	0.025	0.287	0.113	0.910	-0.473	1.269	-0.262	0.793
	Semi-high BM & semi-low risk	-0.034	0.275	-0.178	0.859	-0.047	0.204	-0.261	0.794	-0.074	1.337	-0.102	0.919
	Medium BM & semi-low risk	0.010	0.185	0.075	0.940	0.140	0.169	1.264	0.206	-0.715	0.723	-1.353	0.176
	Semi-low BM & semi-low risk	0.150	0.132	1.484	0.138	0.222	0.112	2.091	0.037	-0.180	0.487	-0.595	0.552
	High BM & medium risk	-0.443	0.106	-0.935	0.350	0.157	0.107	0.719	0.472	-3.746	0.312	-1.388	0.165
	Semi-high BM & medium risk	-0.116	0.471	-0.652	0.514	0.056	0.205	0.363	0.717	-1.161	2.332	-1.594	0.111
	Medium BM & medium risk	0.023	0.173	0.165	0.869	0.148	0.153	1.282	0.200	-0.631	0.660	-0.965	0.335
	Semi-low BM & medium risk	0.087	0.145	0.659	0.510	0.130	0.104	1.003	0.316	-0.022	0.605	-0.048	0.961
	High BM & semi-high risk	-0.740	0.125	-1.938	0.053	-0.291	0.128	-1.537	0.124	-3.306	0.336	-1.544	0.122
	Semi-high BM & semi-high risk	-0.373	0.366	-1.938	0.053	-0.310	0.190	-1.733	0.083	-0.766	1.882	-0.993	0.321
	Medium BM & semi-high risk	-0.165	0.192	-0.887	0.375	-0.226	0.181	-1.289	0.198	0.095	0.748	0.131	0.896
	Semi-low BM & semi-high risk	-0.286	0.205	-1.530	0.126	-0.360	0.192	-1.932	0.053	0.199	0.814	0.310	0.757
	High BM & high risk	-0.685	0.176	-1.493	0.135	-0.127	0.191	-0.408	0.683	-3.571	0.411	-1.532	0.126
	Semi-high BM & high risk	0.002	0.508	0.007	0.995	0.061	0.291	0.175	0.861	-0.209	2.349	-0.148	0.882
	Medium BM & high risk	0.110	0.420	0.299	0.765	0.514	0.376	1.478	0.139	-1.987	1.535	-1.445	0.148
	Semi-low BM & high risk	0.486	0.437	1.060	0.289	0.728	0.379	1.672	0.095	-0.684	1.540	-0.389	0.697
		R^2	0.959			R^2	0.940			R^2	0.972		

Table 6

Regression summary of portfolios independently sorted by BM-ratio and distress risk with three different Fama-Macbeth-regressions run in different business cycles. Standard errors and t-stats corrected for autocorrelation and heteroskedasticity using Newey-West-estimator with lag-parameter of 3.

		2006-2020				Recovery				Recession			
		Estimate	Std. Error	t-stat	Pr(>|t|)	Estimate	Std. Error	t-stat	Pr(>|t|)	Estimate	Std. Error	t-stat	Pr(>|t|)
Independent effects	Reference portfolio (Large & high BM)	0.930	0.382	2.434	0.015	1.272	0.299	4.247	0.000	-0.843	1.885	-0.447	0.655
	Small	0.541	0.159	1.472	0.141	0.190	0.151	0.814	0.416	2.529	0.575	1.236	0.217
	Semi-small	0.241	0.376	1.540	0.124	0.093	0.268	0.891	0.373	1.049	1.544	1.225	0.221
	Medium size	-0.033	0.167	-0.366	0.714	-0.042	0.111	-0.538	0.591	-0.026	0.720	-0.064	0.949
	Semi-large	-0.030	0.080	-0.342	0.733	-0.027	0.082	-0.360	0.719	-0.084	0.239	-0.201	0.841
	Semi-high BM	-0.351	0.082	-2.607	0.009	-0.296	0.077	-2.466	0.014	-0.655	0.253	-1.105	0.269
	Medium BM	-0.144	0.146	-1.075	0.282	-0.181	0.108	-1.773	0.076	0.069	0.719	0.100	0.920
	Semi-low BM	-0.129	0.138	-0.862	0.388	-0.131	0.102	-1.181	0.238	-0.111	0.651	-0.141	0.888
	Low BM	0.125	0.146	0.707	0.479	0.075	0.113	0.564	0.573	0.301	0.672	0.331	0.741
Joint effects	Small & semi-high BM	0.578	0.363	2.301	0.021	0.594	0.288	2.672	0.008	0.557	1.520	0.496	0.620
	Semi-small & semi-high BM	0.136	0.243	0.824	0.410	0.185	0.200	1.237	0.216	-0.078	1.039	-0.110	0.912
	Medium size & semi-high BM	0.229	0.164	1.397	0.162	0.203	0.130	1.424	0.154	0.437	0.771	0.590	0.555
	Semi-large & semi-high BM	0.265	0.183	1.884	0.060	0.342	0.145	2.833	0.005	-0.068	0.878	-0.107	0.915
	Small & medium BM	0.579	0.142	1.990	0.047	0.772	0.119	3.291	0.001	-0.691	0.624	-0.494	0.621
	Semi-small & medium BM	0.273	0.294	1.734	0.083	0.305	0.234	2.090	0.037	0.036	1.424	0.054	0.957
	Medium size & medium BM	0.117	0.179	0.890	0.374	0.143	0.159	1.128	0.259	0.042	0.754	0.083	0.934
	Semi-large & medium BM	0.127	0.123	0.945	0.345	0.202	0.115	1.544	0.122	-0.291	0.463	-0.565	0.572
	Small & semi-low BM	0.852	0.127	2.813	0.005	1.123	0.122	4.091	0.000	-0.779	0.460	-0.616	0.538
	Semi-small & semi-low BM	0.524	0.303	3.192	0.001	0.534	0.296	3.546	0.000	0.429	0.918	0.615	0.539
	Medium size & semi-low BM	0.380	0.187	2.888	0.004	0.323	0.144	2.400	0.016	0.730	0.921	1.653	0.098
	Semi-large & semi-low BM	0.065	0.141	0.478	0.632	0.027	0.139	0.205	0.838	0.263	0.502	0.523	0.601
	Small & low BM	1.607	0.138	4.470	0.000	1.938	0.126	6.085	0.000	-0.372	0.597	-0.241	0.810
	Semi-small & low BM	0.912	0.365	3.623	0.000	0.795	0.357	3.313	0.001	1.677	1.166	1.696	0.090
	Medium size & low BM	0.416	0.272	2.157	0.031	0.315	0.229	1.645	0.100	1.163	1.203	1.739	0.082
	Semi-large & low BM	0.063	0.198	0.378	0.705	0.046	0.173	0.290	0.772	0.288	0.860	0.457	0.647
		R^2	0.968			R^2	0.970			R^2	0.966		