# Do Anomaly Premia Vary over Business Cycles

**Name of student: Madhav Mittal**

**Student ID number: 656054**

Master Thesis: MSc Financial Economics

**Supervisor: Dr. PJPM Versijp**

Erasmus School of Economics

Erasmus University Rotterdam

**Abstract**

The following thesis investigates the variance of financial anomalies over the business cycle and the global financial crisis. The anomalies being studied are the distress risk anomaly and the volatility anomaly. The inquiry first delved into what the strongest form of each anomaly was which yielded results for both the volatility and distress risk anomalies. The inquiry then concerned itself with checking both anomalies for liquidity constraints, checking for variability over business cycles and the global financial crisis and lastly sorted both anomalies over each other. The results of the thesis show that the distress risk anomaly is non linear but sorts cannot be made for the values where the returns peak. The volatility anomaly is strongest when an exact one month holding period is used with weekly returns. The liquidity sorts show that the low volatility anomaly does not hold after liquidity constraints whereas the Distress risk anomaly does persist past liquidity constraints. The results also show that the volatility anomaly responds to the global financial crisis but not to business cycles in the sample. The distress risk anomaly does not respond at all to either of the events being studied. The results thus push the idea that the volatility anomaly is systemic whereas the distress risk anomaly is not. Lastly , double sorting shows that there is indeed an effect of both anomalies on one another with the ideal portfolio being comprised of high volatility high Z score stocks.

# Contents

# 1 Introduction

What determines return? The answer to this question is by definition at the forefront of all financial and equity research. The act of deciding on a concise answer continues to elude authors and inquiries into the matter often raise more questions than they answer. This thesis hopes to contribute to the large collection of literature on the issue. While all research so far admits that returns are not the result of just one factor, many of the discovered factors share similarities. The following thesis also hopes to tease out some of the characteristics of these returns and their relation to the risk borne to achieve them. The best way to find out these characteristics is to study return profiles that go against intuition about the risk return relationship. These abnormalities are also called return anomalies. The first anomaly being studied is the distress anomaly. Introduced by Dichev (1998), the phenomenon shows that companies seen as high risk generate lesser returns than companies that are considered safer. The second anomaly is the Low volatility anomaly. Studied by Heigns(1970), stocks with higher volatility generate lesser returns than companies with low volatility. The distress risk anomaly goes against traditional portfolio theory as firms far from bankruptcy and "too safe" end up out earning their riskier counterparts. This thesis hopes to test the existence of both on the latest dataset and test for how they respond to market cyclicality and the 2008 global financial crisis. The study aims to fortify proof for the existence of both anomalies, followed by testing for the form of the Z-score. Once the existence is verified, the study will move to test their cyclicality as well as their reaction to the global financial crisis. Lastly, the study will test the two anomalies together and see if they influence each other in any way. The study would contribute to current literature in 4 main ways. Firstly, verifying the existence of both anomalies would lend credence to the original theories put forth by those that authored the first papers on the matter. If found to vary significantly over business cycles or economic events (in our case the global financial crisis), we can establish that there is indeed a link between the anomaly premia and underlying economic conditions. This link would lend credence to the idea that anomalies are not simply authors getting "lucky" and that there is indeed economic rationale behind their existence. Secondly, if found to vary over cycles, it would expand our knowledge regarding cyclicality and returns as it would give us definite directions in which cyclicality pushes returns for both the anomalies. Thirdly, as the Z-score anomaly is not as studied as the more popular factors, the study would explain more of its characteristics and show how and why it exists. Lastly, finding a non-linear relationship between Z-score and returns would help contribute to the small pool of literature that exists in the field of non-linear anomalies. The sections for this study are as follows- Section 2 details the literature studied for this paper, section 3 states the questions and hypotheses to be tested. Section 4 and 5 detail the data and methodology used and section 6 and 7 will go over the findings and conclude the text.

## 2 Literature Review

When Harry Markowitz published "Portfolio Selection" in March of 1952, the era of traditional portfolio theory, which was characterized by individual security risk assessments, came to an end. Markowitz along with authors such as Jack. L. Treynor and William Sharpe took the latter half of the 20th century to emphasize the importance of the contribution of a security to an overall portfolio over the risk of the individual security. As a result of this outlook, diversification that kept correlations low among stocks was seen as a superior practice over simple diversification. In that regard Modern Portfolio Theory regards diversification as a means to an end rather than the end destination (Lekovic,2021). The work of Markowitz also lays the ground for the empirical analysis of returns and risk in portfolios as he posited that this relationship was systemic. Andrew Donald Roy who published his work "Safety First and the holding of Assets" two months after Markowitz is credited equally with figuring out this systematic risk return relationship but his work enabled the field further by allowing for negative weightings in his work. In essence, Modern Portfolio Theory cements that risk and return are connected to one another, and you cannot have return without risk. In the next 20 years, opinions emerged contemplating this 'risk'. While the work of Lintner and Sharpe among others pointed to the risk being related to the market (the basis of CAPM), authors such as Lintzenberger and Ramaswamy (1979), Basu(1977) and Banz(1981) started linking the returns of assets to different characteristics. This tangent to Modern Portfolio Theory was not necessarily limiting of the original works as they still posited the same things as Roy and Markowitz but simply stated that the source of this risk was different. While the financial world was still developing these factors, the work of authors around the 1990s gave us tests for these ideas. The works of Fama and French in 1991, 1992, 1988 and 1989 along with Jagadeesh and Titman in 1993 popularised the methodology of creating large portfolios of stocks based on certain characteristics and then testing their returns across historical periods. This is where the work of Roy was important as the portfolios created by these authors consisted of longs and shorts. The essential idea was that going long in a stock rich in a certain factor and short in stocks deficient in the same factor might show us patterns in returns. As both ends of the portfolios were filled with many stocks, their idiosyncratic risk was ignored and more importantly, the stocks would represent an equal mix of other factors apart from the one we are trying to investigate. Also, portfolios constructed this way were perfectly hedged in value and thus the only returns to be made would be from relative movement i.e. one end outperforming the other. These portfolios created were hedged, meaning they posed no apparent risk , and the only difference between the long and short legs was a single factor. As Fama French would explain in their work in the

1993, the Efficient Market Hypothesis still held true as the single factor which was the basis for these sorts was indeed the risk which was being taken. An example of this is the sort based on the size of a company. Smaller companies out earned larger companies and with the rationale being that smaller companies are often more opaque, less liquid and less studied by analysts. This represented a risk loading on small companies that larger companies did not have which earned itself better return.

These results are at the core of anomalies and are currently the subject of great contention. The main opposition to the existence of anomalies is scattered. Arguments span three main ideas. The first part of the opposition posits that returns are explainable by other factors that are not perceived as a risk. The work of Lochstoer and Tetlock(2020) is a great example of this. They show that various anomalies coincide with the cash flows into constituent firms and that cash flow into a firm is less a proxy for risk but rather drives returns on a fundamental level. This fundamental strength however is driven by the upward variance of the cash flows into a company. Another driver for these returns is changes in discount rate as pointed out by Vuolteenaho (2002). Both papers agree that future cash flows are indeed the more dominant contributor to future stock returns. Both papers also agree that a majority of the returns are derived from the variance of cash flows above expectations/forecasts. Vuolteenaho (2002) shows that returns coincide better with cash flows that exceed expectations rather than simply the cash flow. Thus the main contributor of returns as driven by fundamental strength is the variance of cash flows that exceed expectations.

This argument can be explored by looking at the Z score constituent ratio returns as some pose risk whereas some pose fundamental advantages that could drive excessive cash flows. In essence, drawing from the first argument and the traditional notion of risk, stock returns are a mix between the fundamental returns of a firm doing well and a reward for higher risk. The opposite is true where low returns can be due to playing it too safe or a firm genuinely failing to deliver results. This logically makes sense as return is simply the combination of taking a risk and how well it pays off. This is mirrored in Kim(2016) while discussing the Z-score anomaly.

The second argument is that anomalies are the result of market inefficiencies. This however isn't as much an argument against the anomaly but rather its viability. Most papers that address this criticism show that anomalies are only theoretically possible as the issues of liquidity and transaction costs of trading on these anomalies far exceed the returns generated from said anomaly.

This fact is echoed very strongly by Jensens's 1978 paper who describes prices as reflecting information till the point where the marginal cost of acting on the information is greater than the benefit. This form of market efficiency is also endorsed by Fama in his 1991 paper. In that regard, this argument is still compatible with Fama and French's work. Regardless, this can be tackled by proxying for liquidity and checking the anomalies.

The third argument is biases related to publishing and p-hacking. Hou Xue and Zhang (2020) are perhaps at the forefront of this argument as they found it hard to replicate a majority of anomalies in a different time period with more international data. This paper also puts an emphasis on the replication of anomalies. The authors express concern that the publication of anomalies in papers is incredibly biased as there is often no publication of non-result studies and replications are hard to get published. By attempting to replicate two anomalies, one often studied and the other nearly abandoned, this study hopes to lend credence or take it away from the original authors.

The distress risk and volatility anomalies are counter intuitive since both of them posit that there is money to be made in betting in the safer ends of these spectrums.This goes directly against the ideas put forth by Fama and French as both low risk of distress firms and low volatility firms pose lesser risk to investors.

Distress risk was first used to sort portfolios all the way back in the 1993 Fama French paper where they calculated the spread between corporate and government bonds for the exact same tranche and time. This spread explained stock returns but as an extension of the size anomaly with the rationale that smaller companies are more likely to fail. More importantly, the original paper takes the spread as a proxy or probability of default but there is no critical point at which a firm would go from healthy to distressed.

Thus the measure used to predict bankruptcy risk must be accurate and predictive of distress.

The approach I will be taking to measure risk is Altman's Z-score. Z-score as it will be used in this study was invented by Edward Altman in his 1968 paper. There are four main questions that must be answered to justify the use of the Z-score.

1. Why ratio analysis is the right choice for our purpose.
2. Why the ratios chosen in the Z-score are the right ratios to predict bankruptcy.
3. What the weights of each of these ratios should be, and;
4. Is this ratio still a good descriptor of bankruptcy risk.

To answer question one, it makes sense to firstly compare different methodologies. There are two main schools of predicting bankruptcy, these are parametric and non-parametric methods (Min, Jeong (2009)). Parametric methods include Regression analysis, Hazard Analysis and Multiple discriminant analysis. The other school entails use of neural networks, clustering and decision tree forests. The former is prefered for two main reasons. Firstly, parametric methods demand lesser computational power (Almaskati et.al 2021). The second is the issue of overfitting of data in the case of non parametric methods.

Horak et.al (2020) for example, compile a list of 22 variables that they used in a model multiple layers deep and while they did find better results with their model and higher out of sample results for the same industry over time, they admitted that the fitting meant that the model cannot be exported to other settings. Pisula (2020) finds similar problems with their model being much weaker in different industries. The authors describe this as being an issue of how specific variables can be in an industry. Horak et.al(2020) for example uses the riskiness of managerial decisions into account in their neural network and admits that a variable as specific as this can be vastly different for different industries. The authors also posit that these differences are larger for small companies than they are for larger more established firms. In that regard, ratio analysis while losing out predictive power by one or two percentage points is more comparable across sectors and in a larger sample will do better (Masten, Masten (2007)). Another advantage of the methodology of simpler ratios is the matter of accessibility. As we study the efficient market hypothesis and the valuation of stocks by the market, it makes sense to use a method that returns would be more responsive to. As Ohlson(1980) found that stock movements were a lot more responsive to information right when reports were released, it makes sense to choose a methodology that would accurately represent the rate of dissemination of information into the stock market. Neural networks might be better predictors of bankruptcy, but ratio analysis would more accurately represent market sentiment regarding the potential failure of a stock as ratio information is available a lot easier wheras a lot of neural network models use information found ex post.(Karas et.al 2014).

To answer questions 2 and 3, it makes sense to look at the methodology and weights of the actual Z-score. Altman derived the score through an a priori clustering technique known as Multiple Discriminant Analysis or MDA. MDA works by considering quantifiable characteristics of individual observations and uses them to demarcate a line between two or more groups. It works by creating r-1 lines between groups where r is the number of groups. In the case of distressed and non-distressed firms, there as the need of only one line and this line was thus the linear combination of 5 different ratios-

$Z = 0.012X1 + 0.014X2 + 0.033X3 + 0.006X4 + 0.999X5$
Where-

- X1 = working capital/total assets
- X2 = retained earnings/total assets
- X3 = earnings before interest and taxes/total assets
- X4 = market value of equity/book value of total debt
- X5 = sales/total assets

The ratios chosen in this line are derived mainly from prior literature and are chosen by the author based on how well these ratios combined can predict bankruptcy. This method also meant that individual predictive profiles were ignored in favor of their contribution to the line. As a result, some significant variables were left out as their effect on the predictive power fell when used in tandem with the rest.

For Answering question 2, describing each variable might do a better job of showing why it exists in the list.

The Working capital ratio (X1) represents how well the firm can fund current activities based on the size of the firm. A higher working capital ratio represents a high liquidity for the firm. A constraint of liquidity is both an indicator of previous financial failures as well as future financial distress (Altman, 1968). Firms lose liquidity due to lessening current assets against current liabilities which can then lead to financing becoming more expensive. Liquidity constraints also lead to a higher probability that debt will not be repaid which is the definition of financial distress (Vincent et.al 2019). Thus liquidity is perceived as being good for the firm as Rahman Iqbal and Nadeem(2019) and Anton and Nucu(2021) found positive effects of the ratio on a company's returns. Wang(2019) also sorted portfolios based on the factor as a part of their attempt to find a cash conversion cycle spread. The ratio in Wang's analysis was the most important in explaining the Cash Conversion Cycle. The sorts Wang found are indeed significant with deciles providing a positive profit with higher liquidity firms doing better. This again points towards the Working Capital Ratio being a good sign of liquidity and important in determining the success of a company. More importantly, it points towards safer stocks having more return.

The ratio of Retained Earnings to Total assets (X2) is like the CMA factor put forth by Fama and French in their 5-factor model (2015). It represents how much assets in the future will grow. Altman also sees this ratio as inherently representing the age of a company, as companies often fail in early stages according to the author. Low retained earnings thus could come from a firm that has been recently founded, and for firms in later stages represents lower earnings. Both of these are hazardous for a firm which boosts how well it predicts bankruptcy. While the Authors found CMA to be positive- meaning higher retained earnings mean lesser returns, O'Donnovan (2021) found results like theirs in their work using the Retained Earnings to Total assets ratio, higher returns for lesser retained earnings. This points towards safer stocks being less profitable.

EBIT by total assets (X3) or Basic Earning Power Ratio falls into the same pool as X2 as it often represents very similar earning power. However, as Baker Hoeryer and Wurgler (2016) point out, the two vary based on firm Leverage. In our case, we would expect the two to behave in the same manner.

Equity by Debt (X4) as described by Altman shows the buffer that a company has before it is encumbered by too much debt. It represents how much value a firm can lose before bankruptcy is declared. A higher Equity to Debt ratio calculated as such is good for a firm hoping to reduce the risk of insolvency but does not represent higher returns as lower leverage leads to lower returns according to Altman (1968).

Lastly, Sales by total Assets (X5) or Asset Turnover is how effectively the company can generate sales. It is an efficiency ratio much like every part of the cash conversion cycle. Vedy and Santoso (2022) find that it positively contributes to the firm's growth and find a positive relationship with equity returns.

We now find that Z-score is made up of 5 ratios that differ in their relationship with return. The line drawn by these 5 factors with these factor weights creates the best discernible line between firms that are distressed and not distressed. The author states that the deficiency of any one of these factors is not often enough to be distressed and the regression line obtained would fail to account for that fact. We now move to answering questions 3 and 4 together.

Cindik and Arnutlulu (2021) point out that while the weights themselves are derived from an optimization of the problem, they do not lack reasoning. The reasoning for the ratio weightages is their relative importance to the overall solvency of a company. With the rankings laid out as they are-

- X5 - Sales/ Total assets is weighted as the most important as written in Altman et.al (2007) with sales generation being cited as the main fundamental strength of a business.
- X3 - EBIT/Total Assets is written as the second most important in the list. Ohlson (1980) writes that operational profitability is also an important predictor but must still be weighted lower than sales generation as volume of sales is more important than profitability. Ohlson also writes that high sales volume is indicative of profitability as a firm with low to no profitability is unlikely to carry out a large number of sales.
- X2 - Retained Earnings/Total Assets is ranked third in importance in the ratio. Altman (1977) writes that Retained earnings as a proxy for both age and cumulative earnings is still important to a firm but not nearly as important as the previous measures since it can vary more greatly based on the discretion of leadership.
- X1 - Working Capital/Total assets is ranked second to last since liquidity is only a short term solution to solvency in case of the failure of the rest of the three ratios (Deakin (1972)). It is however still more important than our last measure;
- X4 - Equity/Total Liabilities is indicative of market sentiment as well as a secondary buffer to insolvency that firms possess. Beaver(1966) writes that this ratio is very

lowly held in importance since the secondary buffer that it provides is very rarely used by publicly listed firms but is nonetheless important to the solvency of a firm.

Establishing this hierarchy, we can see how papers maitaining these rankings but with slightly different weights found predictive results. With various weightages possible, the original Z-score still finds the most robust line in the USA stock universe. The predictive power of this line has varied based on the setting and the time of the data. Gordini (2014) found the line to be correct for 84.4% of companies in a set of 3100 companies over 30 years from 1980 to 2010. Zelenkov et al., (2017) found the model to be predictive for 93.4% of the 914 companies studied in Russia. Pantoja-Aguilar (2022) found that the line was predictive for 86% of companies that they studied. Cindik and Artmlulu in 2021 found an accuracy of 85% in the companies that they studied in Turkey. Altman's own rerun of the model in 2000 in the USA was correct 95% of the time for companies studied in their sample. While all of these studies found some variation of the score to have more predictability outside of the USA, the original score was still the most accurate in the case of companies in America. More importantly, the improvements of other models were in the prediction of solvent companies whereas the Z score performed equal or better in terms of the prediction of bankruptcies outside of the USA. This provides us with the best criteria to create our long and short portfolios with longs in high Z score (relatively safer) companies and shorts in low Z score (relatively riskier) companies. With that in regard, it makes sense to use the Z-score as Altman wrote it as our sample is going to be American companies.

The preliminary paper putting this score into the process of sorting portfolios is the work of Dichev (1998) who concluded that the anomaly was simply the case of mispricing due to distressed firms being less liquid. This idea was mainly driven by the fact that distressed firms are often far less liquid than non-distressed firms and can be compared to small firms. This failed to explain however why the returns Dichev found were hump shaped. The author states that the returns were in a U shape. This meant that companies with moderate Z scores found excessive returns wheras firms with too high or too low Z Scores found lower returns. This also explains why the author found it hard to explain this as a prt of the size anomaly as the size anomaly is linear in nature and size has a negative but linear relationship with returns. The paper that succeeded the most to explain this was the work of Kim (2016). The author explains the hump as the reconciliation of the works of Fama and French who say that distress risk is positively correlated with returns and the works of Campbell, Hilscher and Szilagyi (2008) who found the opposite relationship. These two effects are essentially explained by the Assumption we reached in the previous section that higher returns are a mix between the fundamentals of a company and the reward for higher risk. Using an accounting ratio like the Z score for these default probability estimations is what causes the hump shape due to a low Z score representing a high distress risk whereas a high Z score represents good company fundamentals (Kim 2016). Kim's biggest contribution in this paper however is showing that there is a peak in the hump (firms with Z scores around 5 have the highest returns). This can be seen as the optimal mix of risk and fundamnetal strength as described by Kim, and is the point where the risk effect (outlined by Fama and French) is overtaken by the Fundamental effect(Campbell et.al (2008) and Lochstoer and Tetlock (2020))

In that regard, we have a structural form that we can test for. Finding the apex of this graph would thus give us a stronger anomaly to test than simple linear sorts. More importantly, finding a similar apex across companies might allow us to point out the contribution of risk and fundamentals to return. More importantly, The DEF factor represents purely risk by design. This is indicated by Fama and French's explaination that a higher risk of bankruptcy is the only factor that explains the spread between government and corporate bonds with the same charecteristics. The Z-score however constitutes 5 different ratios that represent both risk and fundamental return generating capabilities. In that regard it is better suited to explore the question raised by Lochschoter and Tetlock.

The second anomaly studied is the volatility anomaly. This anomaly was first investigated in the 1970s with the work of Robert Haughen James Heighns who found results were more robust for low volatility portfolios in the 50 years they studied. More recently, there were similar results found by Li, Garcia and Feijoo (2014), Clarke Silva and Thorley (2010), Beijer (2015) and Blitz and Van Vliet (2007). While all of these papers regard the anomaly as existing, Li, Garcia and Feijoo put forth that the Anomaly is reduced by the limits to arbitrage which coincides with the argument against anomalies as a whole. The authors however do find cyclicality in the anomaly. The reason for the low volatility anomaly is mainly broken down into 3 different arguments. These are mispricing, systematic risk and liquidity. Ang (2009) finds the volatility anomaly coincides in different countries branding it a systematic risk rather than simply something diversifiable. This would be most in line with the argument that the volatility is cyclical as cyclicality cannot necessarily be diversified away. Thus, finding a variance across economic factors would put this risk as being a part of the systematic risk profile of a stock.

Reading on varying returns from anomalies, there are conflicting opinions on the matter but volatility seems to be the most prone to cyclicality. This is also in line with Ang (2009) that claims the anomaly to be systemic. As stated earlier, the authors on volatility anomalies do indeed believe that these are cyclical. Jacobsen Mamun and Visaltanachoti (2005) tried similar methodologies with the Size and value anomalies but instead of using cycle data they tested for the "Halloween effect". The papers created sorts and then tested the occuring returns against a dummy for the months considered in the Halloween effect to see if there is any variance. The results proved to be inconclusive and could not prove a difference. The most that we know however on cyclicality comes from Blitz and Van Vliet (2007) who found a downturn strengthened the anomaly and thus managers working on that anomaly were mostly expected to outperform during downturns.

# 3 Research Question

With that being said, the proposed study proposes three major questions to be answered-
1. Do the Z-score anomalies and Volatility anomalies still exist controlling for transaction costs?
2. Do these anomalies show variance over time and over different phases of the business cycle?
3. Do the anomalies interact with one another?

Answering Question 1 will help us restate what prior authors have stated. This would also include answering what the form of the Z-score anomaly is which promises to have interesting results and would help us decode what exactly drives the returns. This would be a direct inquiry into Criticism 1 discussed earlier. This process would also include using proxies for liquidity in working out what anomalies would survive the liquidity criticism.

Answering question 2 would help us provide an insight into both the nature of cyclicality and the anomalies themselves, as well as whether they are systematic or not.

Answering question 3 is merely a case of exploring both anomalies further as the Z-score portfolios are often sorted with size but not volatility.

# 4 Data

## 4.1 Volatility

The universe of stocks that will be used for the study is sotkc market data in the USA. The companies considered come from the NYSE, NASDAQ and AMEX and only common stock is considered. As a result, the data is sourced from the CRSP return database using the monthly stock files. Stocks below 5 dollars will be dropped as well as financial firms. To measure the volatility, I will take the same approach as in the paper of Blitz and Van Vliet (2007). I propose to calculate their past 3-year and 1 year volatility using weekly volatility. The volatilities will be calculated by aggregating daily returns into weekly returns, calculating volatilities for each of the weeks using a rolling window of the previous 52 and 156 weeks. To make the data monthly, the volatility of the last week in every month is assigned to that month, thus making only one volatility value per month. Because only the last week of volatility is used, returns data for the entire month does not represent the difference between two monthly measurements of volatility as it is offset by 1 week each time. For example, measurements taken in the last week of June and the last week of July cannot be compared using monthly return data of July as it would skip the last week of June and include the last week of July. Thus, the return data used for volatility will be "weekly returns" which is the returns data used to calculate the volatility in the first place. Thus, comparing the weekly returns of the last week of two months gives us a better picture and should be more representative. This also fixes our problem of trying to make a month-long holding period as we can simply take a stock and look at the weekly return from the last week of the next month. To be transparent however, returns for the month are also used for the volatility portfolios and weekly returns are also used for Z score portfolios.

### 4.2 Altman's Z score

Altman arrives at the Z score through multiple discriminant analysis and out of each category of financial ratios chose the best one. The score can be calculated as a simple sum of each ratio multiplied by the weights specified in the equation. The Z score would be calculated via the Compustat merged database using fundamentals for calculations.
The original Z score was of the form-
Z = 0.012X1 + 0.014X2 + 0.033X3 + 0.006X4 + 0.999X5
Where-

- X1 = working capital/total assets (WCAPQ/ATQ)
- X2 = retained earnings/total assets (REQ/ATQ)
- X3 = earnings before interest and taxes/total assets ((REVTQ-COGSQ-XOPRY)/ATQ)
- X4 = market value of equity/book value of total debt ((CSHOQ*PRCCQ)/(DLTTQ+LCTQ)
- X5 = sales/total assets (REVTQ/ATQ)

With each being the indicator for liquidity, profitability, efficiency, solvency and sales generation respectively. The calculation of the Z-score is simply to calculate each of the ratios for a company and then add them up with these weights.

The score was for publicly traded companies and will be used for the entire universe of stocks. The available data in the Compustat merged database for the same variables is only available from 1961 to 2020 and this entire range of data will be used.

### 4.3 Business Cycle Data

Data for business cycles is available via NBER starts in 1857 and is available till April 2020. The data allows us to classify all the business cycles as troughs, peaks, expansion and contraction with periods for each of these. However, only the relevant data from

1960 to 2020 will be used.

# 5 Methodology

The results presented in section 6 are the result of methodologies constructed to conduct an inquiry into the questions and criticisms raised earlier in this text. This section is an outline of how these methodologies were constructed and what their purposes are in the context of our questions. The methodology is concerned with first the construction of the factors, establishing a form for the Z-score anomaly, single portfolio sorts, Fama Macbeth Regressions and double sorts of the portfolios, each factor individually with a bid ask spread to proxy for illiquidity and a second double sort between volatility and Z score to estimate the effects they have on each other.

## 5.1 Determining the form of the Z Score Anomaly

The main issue with the Z-score is that the relationship with returns has been found by previous authors to be nonlinear. While a nonlinear anomaly is not alien, it does make it harder to decide on portfolio sorts. The main aim is to find an apex that is within the data range as well as consistent across datasets. Thus, if the regression run between returns and the Z-score is indeed nonlinear, then we split the data set into 2 30 year periods. This is to test for whether or not there is a consistent peak across the 60 years of data. The regression will first be run only for the first period, from 1960 to 1990. The generated equation will then yield a hypothetical maximum. If this maximum is within the data range, it will then be taken as Z* and portfolio sorts will be run on the second period to see if there is indeed a consistent apex to this curve. This is however dependent on finding the apex within a reasonable data range. If the apex is too far out, then the relationship at least for our intentions is no longer quadratic.

## 5.2 Portfolio Sorts

The portfolios will be created based on Volatility and Z score, both will be rebalanced monthly with value weighted returns as well as equal weighted returns, used to see whether the long in top and short in the bottom (vice versa for volatility) will provide positive returns. These will be done for the sample as well as separately for expansion/contractions post which the mean returns will be tested for differences.

$H_0$: Firm with high Z scores/volatility do not have different returns than firms with low Z scores/Volatility
$H_a$: Firm with high Z scores/volatility have different returns than firms with low Z scores/Volatility

$H_0$- r(P10-P1) $= 0$
$H_a$- r(P10-P1) $\neq 0$

With P10 being the highest Z and Volatility returns and P1 being the lowest Z and volatility returns. The results of these tests will first help us establish whether these anomalies exist in our dataset.

### 5.3 Testing Against Component Ratio Returns

Portfolio sorts are run for all of the component ratios of the Z-score ri,t(Z), generating a series of returns that represent each of the risks. The returns will then be tested against the original Z-score returns to show which of the component ratios is most important to this score.

$$r_{i,t}(Z) = \beta_1 \cdot r_{i,t}(X_1) + \beta_2 \cdot r_{i,t}(X_2) + \beta_3 \cdot r_{i,t}(X_3) + \beta_4 \cdot r_{i,t}(X_4) + \beta_5 \cdot r_{i,t}(X_5) + \varepsilon_{i,t}$$

$$\begin{aligned} H_0 &: \text{ The constituent ratio returns have no effect on anomaly premia.} \\ H_a &: \text{ The constituent ratio returns have an effect on anomaly premia.} \\ H_0 &: \ \beta_1 = \beta_2 = \beta_3 = 0 \\ H_a &: \text{ At least one of the Betas } \neq 0 \end{aligned}$$

This section has the potential to be interesting as it may help explain the humped relationship in the anomaly.

### 5.4 Fama Macbeth Regressions

Each of the return cross sections for all three variables are tested in a Fama Macbeth style regression for both the business cycle and a dummy variable for the global financial crisis.

$$r_{i,t} = \alpha + \beta_1 \cdot \text{Vol}_{i,t-1} + \beta_2 \cdot \text{Dummy Contraction} + \beta_3 \cdot (\text{Vol}_{i,t-1} \times \text{Dummy Contraction}) + \varepsilon_{i,t} \tag{1}$$

$$\begin{aligned} H_0 &: \text{ Volatility of a stock and the phase of the business cycle has no effect on anomaly premia} \\ H_a &: \text{ The phase of the business cycle influences the anomaly premia} \\ H_0 &: \ \beta_1 = \beta_2 = \beta_3 = 0 \\ H_a &: \text{ At least one of the Betas } \neq 0 \end{aligned}$$

Where Beta 1 is the indicator of the anomaly and Beta 2 and Beta 3 are indicative of the return changing over business cycles. A similar regression is run with Z with the mentioned dummies. However, for the Z score, there will also be a test on the Z Squared. This regression is the main one used to try and find the apex

Another form of the regression is run for effects before and after the 2008 GFC. This regression is likewise be run for Z scores

$$r_{i,t} = \alpha + \beta_1 \cdot \mathrm{Vol}_{i,t-1} + \beta_2 \cdot \text{Dummy 2008} + \beta_3 \cdot (\mathrm{Vol}_{i,t-1} \times \text{Dummy 2008}) + \varepsilon_{i,t}$$

$$\begin{aligned} H_0: &\quad \text{The anomaly premia do not change after the 2008 GFC.} \\ H_a: &\quad \text{The anomaly premia do change after the 2008 GFC.} \\ H_0: &\quad \beta_1 = \beta_2 = \beta_3 = 0 \\ H_a: &\quad \text{At least one of the Betas } \neq 0 \end{aligned}$$

Where $\beta_2$ and $\beta_3$ are indicators of the effect of the 2008 crisis on returns. The time period for this regression is from 1998 to 2018 to account for an equal horizon on both sides of the crisis.

Ideally, I believe that as the sentiment changed post 2008 to a more conservative risk avoiding one, the low volatility anomaly should be lower in this period. In essence, the Fama Macbeth regressions should tell us about
1. The shape of the Z score anomaly
2. The effect of the global financial crisis and contractions on both anomalies.

### *5.4.1 Fama Macbeth Rgressions Versus Regressing Returns*

The question of testing whether an anomaly varies over time can mean two different things. The first is that the spread between the highest and lowest deciles is higher and secondly, with the spread between deciles remaining the same does each "unit" of our factor lead to more returns.

These two are in fact different things and both equally important in answering this question. The first one is answered by regressing the returns from the portfolio sorts against the GFC and contraction dummies. This will tell us what the absolute change was in the returns without saying anything regarding the change in the spread of deciles themselves. Hence, we need to investigate further. The Fama Macbeth regressions on the other hand, are a lot more telling. Creating an interaction effect for the dummies and the factors, we can regress the entire sample and see whether the Beta of any of the factors increases or decreases when the dummies are one. Finding this out can help us verify whether the different returns are derived from a larger spread or each factor being more potent.

### 5.5 Double Sorts

The factors are then double sorted, first individually with bid ask spreads to look at whether the returns are driven by illiquid stocks, and then volatility will be sorted with Z. This should indicate a relationship between the two as Kim (2016) did a similar procedure between Z score and size allowing him to show the effect of Z score on the size anomaly.

The sorts will be split for the effect of Z score anomaly on volatility anomaly and the effect of volatility anomaly on Z score anomaly.

For the effect of Z score
$H_0$- The Z score has no effect on the size of the volatility anomaly Premia
$H_a$- The Z score has an effect on the size of the volatility anomaly Premia
$H_0$- ri,t(V1,Z10)- ri,t(V1,Z1) = 0 and ri,t(V10,Z10)- ri,t(V10,Z1)= 0
$H_a$- At least one of the differences $\neq 0$

For the effect of Volatility
$H_0$- The Volatility has no effect on the size of the Z score anomaly Premia
$H_a$- The Volatility has an effect on the size of the Z score anomaly Premia
$H_0$- ri,t(V1,Z1)- ri,t(V10,Z1) = 0 and ri,t(V1,Z10)- ri,t(V10,Z10)= 0
$H_a$- At least one of the differences $\neq 0$

## 6 Results

### 6.1 Descriptive Statistics

In order to parse the results of the inquiry conducted better it is of the essence to understand the data being used better. The data sourced from WRDS comprises the universe of American Public stocks. The initial count of companies considered excluded firms listed in smaller exchanges as well as financial firms as the debt profile of such firms is vastly different to that of the rest of the universe. The starting number of firms considered from the initial data provision by WRDS was close to 26,000 firms. This number was then reduced to a little over 17,000 firms after adjusting for inconsistencies as well as duplicate and missing values. Firms that closed down during the sample period were also included in the study as with an average listed vintage of 9.33 years, it cannot be reasonably assumed that each firm would survive from the beginning to the end of the period.

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| RET | 1,898,083 | 0.012 | 0.194 | -0.994 | 24 |
| Z score | 1,898,083 | 0.331 | 0.263 | -2.231 | 36.065 |
| X1 | 1,898,083 | 0.265 | 0.263 | -35.853 | 2.157 |
| X2 | 1,898,083 | -0.476 | 2.849 | -99.259 | 55.26 |
| X3 | 1,898,083 | -0.236 | 0.389 | -29.887 | 2.54 |
| X4 | 1,898,083 | 6.365 | 11.093 | 0 | 100 |
| X5 | 1,898,083 | 0.305 | 0.266 | -1.751 | 36.915 |
| vol 1year | 1,898,083 | 0.077 | 0.051 | 0 | 5.003 |
| vol 3year | 1,898,083 | 0.078 | 0.046 | 0 | 5.003 |

Table 1: Summary of Descriptive Statistics for Variables

Table 1 displays the characteristics of the main variables used in the study. These include return which was used in both daily and monthly frequencies. The one tabulated below is the monthly frequency. Also included are the variables used for both sorts and the regressions in the later stages of the study. One of the main things to be drawn from the table however is the wide ranges of X1-X5 included in the study. The main reason that these were not trimmed or winsorized is due to the original specification. Altman mentions in his study that the ratios for smaller firms can often reach extents as large as these but as long as the final Z score is trimmed and is within reasonable limits; the range is acceptable.While Altman does not specify what limit counts as reasonable, papers further using the score (Clarke et.al (2010)) uses a maximum range of +100. This is thus the same cutoff point used in this study. The measures for X1 to X5 were also similarly trimmed which led reasonable values.Notably, X4 had . Coming lastly to the volatility measures, it makes sense that the minimum would be zero as negative volatilities are not feasible.

The measures for price and liquidity are listed in table 2 below. These measures also are obtained after our discounting of penny stocks.

| Variable | Obs | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|
| PRC | 1,921,322 | 19.575 | 43.912 | 1.02 | 4,505 |
| SPREAD | 244,699 | 0.444 | 0.873 | 0.001 | 140.95 |

Table 2: Descriptive Statistics - Price

The mean return over the 50 years as given by WRDS is in log form. As a result, the data is fairly stable and free of any trend or seasonality. With the large dataset of companies being studied this result is not surprising either as the diversification would lead to a mean around zero

### 6.2 Time Series Graphs

Graphs 1.1 and 1.2 of the appendix follow the historical value of volatilities in the sample. The volatility calculated for both 1 and 3 years is indicative of market conditions as the spikes are around periods where volatility is expected to rise (namely the 2008 global financial crisis and the Covid-19 pandemic). The 3-year volatility which was constructed across the previous 156 weeks has a much lower amplitude than the 1-year volatility as the number of weeks in the rolling window allows for a relatively smoother graph.

Graph 1.3 displays the historical value of working capital by total assets. X1 which is the working capital turnover ratio is a short-term liquidity measure. As expected, it began to turn downwards around the 2020 Pandemic as the halting of businesses affected how well firms could hold onto liquidity. Bellucci et.al (2022) show that this was indeed the case due to two main reasons, Lower sales reducing liquidity in most firms and the lack of liquidity in the system exacerbating the problem.

Graph 1.4 tracks Retained earnings to Total assets(X2). As a measure for future growth, the steady decline from the 1980s onwards is indicative of firms developing less enamoured with the measure. In fact, it might be argued that lower retained earnings might indeed be a measure of growth as average RE is just below 0. Jenkins (2003)put forth this argument as well, that Retained earnings are also due to large amounts of acquisition and investment by a firm which leads to future growth. The authors do still posit that this is a risky move in most cases, and that firms with positive Retained earnings are still a safer bet. This thus still represents a risk that we are willing to study. This also ties this back to the comparison to the CMA factor from Fama French.

Graph 1.5; X3(EBIT over total assets) starts out negative in this sample moving closer to zero near the end. While also negative overall, the opposite pattern to Retained earnings confirms that as firms started to make more, they began investing more back into themselves. However, the average is always slightly below zero. As X3 is taken as a percentage in the formula, the real semi deviation below zero is closer to -.02% of total assets, making more intuitive sense.

Graph 1.6 shows that Equity by debt(X4) stays relatively stable till the late 2010s and going into 2020. Abraham Cortina and Schmuckler(2020) show that most of the debt taken on post 2008 has been in developing economies and thus the American market had fairly stable equity to debt ratios.

Graph 1.7 shows that the sales to total assets ratio has been declining consistently past the 1990s. Akcigit and Ates (2019) put the fault of this on reducing dynamism and that higher concentration of established firms is making margins reduce and bringing down average profitability ratios.

### 6.3 Portfolio Sorts

Portfolio sorts still exist as the main way to check for the existence of an anomaly. While sorts can be made as far ranging as percentiles to quintiles, this study utilizes deciles to carry out the sorts. The sorts were conducted with 8 variables: prior 1 year volatility, prior 3 year volatility, prior months Z score, prior months X1, X2, X3, X4,X5. Assuming that the portfolios were held for a month, the difference in mean return between the highest and lowest deciles (deciles 10-1) was regressed first on their own to show if there is indeed a difference in the returns of each. Also used in these regressions were the mean excess returns over the risk-free rate as well as the weekly return calculated in the measuring of volatility.

These differences were then run against the Fama French 3 factors: The Excess Market Return, Small Minus Big(SMB) and High minus Low(HML). In the next regression the variables added were Conservative Minus Aggressive(CMA), Robust Minus Weak(RMW) and Carhart Momentum (Up Minus Down (UMD). As written earlier, a positive alpha (intercept) in each of these cases is indicative of the existence of an anomaly. An Alpha in case of the FF3 and FF5 regressions is indicative that the factor returns as constructed cannot be captured by these already existing anomalies. Lastly, for volatility we would want the alpha to be negative since we use highest volatility minus lowest volatility.

#### *6.3.1 Volatility 1 year*

The Volatility factor equal weighted returns (Table 3) produce no anomaly in the sample. These results are explored very similarly by Mulder (2020) using equal weighted returns and decile sorts. The studied time period (1963-2019) in that study makes it suitably comparable. The study showed no return in a low minus high decile sort. Sullivan and Garcia Feijo (2014) find similar results past 1990 finding similar returns from low and high volatility portfolios.

Regardless, the FF 3 factors are all influential on the returns and can explain a majority of the variation of return. Using the weekly Returns measure which is more accurate of a measure for a 1 month holding period (Table 4) showed that the anomaly did indeed exist and that it is significant with a .4% or .5% alpha when controlling for 3 or 5 factor risk adjustments as seen by the negative and significant intercepts.

Table 3: **Volatility 1 Year**

| | Vol 1 Year | Vol 1 Year | Vol 1 Year | Vol 1 Year |
|---|---|---|---|---|
| **Excess Return on the Market** | 0.360** (0.083) | 0.272** (0.087) | 0.234* (0.094) | |
| **Small-Minus-Big Return** | | 0.258* (0.130) | 0.151 (0.138) | |
| **High-Minus-Low Return** | | -0.316* (0.130) | -0.218 (0.181) | |
| **Robust Minus Weak Return** | | | -0.434* (0.185) | |
| **Conservative Minus Aggressive Return** | | | -0.151 (0.282) | |
| **Momentum** | | | 0.012 (0.091) | |
| **Intercept** | 0.005 (0.004) | 0.003 (0.004) | 0.004 (0.004) | 0.005 (0.004) |
| **Number of observations** | 595 | 595 | 595 | 595 |

*Note:* ** $p < .01$, * $p < .05$

### *6.3.2 Volatility 3 year*

Much like the previous section, 3-year volatility (Table 5) also shows positive but insignificant trends with most of the variation in returns explained away by the 3 factor Fama and French returns. The calculations with weekly Returns however (Table 6) show that the anomaly exists when one months exact holding period is used. This is also seen in the negative and significant intercepts.

### *6.3.3 Z Score*

The Z score anomaly shows indeed a positive alpha in the cases of 3 and 5 factor Fama French regressions (Table 7) of 2% per month. The alpha is unchanging in the case of the weekly return measurement (Table 8). The results agree with those of Kim (2016) Dichev(1998) and Campbell et.al (2008). All of these results find in some capacity that lower distress risk is associated with a higher return. This goes against our understanding of the risk return relationship and in that regard is also in the same vein as the low volatility anomaly. The reasoning behind this result has many theories. The first put forth by Campbell et.al (2008). Theory 1 states that the result is only an in-sample anomaly and that it cannot be replicated. This theory can be tested by looking outside the sample the authors claim this is true for (they test till 2000). Appendix table 16 is the same regressions but only post 2000. The results are similar and the alphas stronger, hence this theory does not seem to explain the reasoning for this anomaly. The alpha in the case of the 5 factor FF model is 2.9% per month. The second theory is put forth by Dichev that the stocks that are distressed are small and less liquid. This is interesting

Table 4: **Volatility 1 Year with Weekly Returns**

| | Vol 1 year | Vol 1 year | Vol 1 year | Vol 1 year |
|---|---|---|---|---|
| **Excess Return on the Market** | | -0.242** | -0.185** | -0.158* |
| | | (0.027) | (0.028) | (0.029) |
| **Small-Minus-Big Return** | | | -0.266** | -0.167** |
| | | | (0.041) | (0.042) |
| **High-Minus-Low Return** | | | | 0.082* |
| | | | | (0.041) |
| **Robust Minus Weak Return** | | | | 0.388** |
| | | | | (0.057) |
| **Conservative Minus Aggressive Return** | | | | 0.026 |
| | | | | (0.087) |
| **Momentum** | | | | 0.010 |
| | | | | (0.028) |
| **Intercept** | -0.005** | -0.004** | -0.004** | -0.005** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| **Number of Observations** | 596 | 596 | 596 | 596 |

*Note:* ** $p < .01$, * $p < .05$

Table 5: **Volatility 3 Year**

| | Volatility 3 Year | Volatility 3 Year | Volatility 3 Year | Volatility 3 Year |
|---|---|---|---|---|
| **Excess Return on the Market** | | 0.394** (0.082) | 0.298** (0.087) | 0.261** (0.094) |
| **Small-Minus-Big Return** | | | 0.262* (0.129) | 0.173 (0.137) |
| **High-Minus-Low Return** | | | -0.356** (0.129) | -0.270 (0.180) |
| **Robust Minus Weak Return** | | | | -0.362* (0.184) |
| **Conservative Minus Aggressive Return** | | | | -0.152 (0.282) |
| **Momentum** | | | | -0.007 (0.091) |
| **Intercept** | 0.003 (0.004) | 0.001 (0.004) | 0.002 (0.004) | 0.004 (0.004) |
| **Number of observations** | 595 | 595 | 595 | 595 |

*Note:* ** $p < .01$, * $p < .05$

as Fama French also found a good amount of shared variation between size and distress risk. But Dichev claims that this is due to illiquidity while the Fama French paper claims that small stocks are more likely to be distressed. The direction of causality is debated on and while it is hard to prove one or the other, we can eliminate the liquidity claim by testing for it. This is done in a later section

The results of the portfolio sorts thus can be interpreted as being strong for the case of the Z score and weak at best in the case of low Volatility. We can reject the null in both cases however for volatility we need to use a specific measure of return. This discrepancy comes down to the way the factor is constructed. Using only the last week of each month, we are looking at a period from the last week of every month to the next. This is not the case for the Z score as the quarterly data is applied to each of the months, thus a

Table 6: **Volatility 3 Year with Weekly Returns**

| | **Vol 3 year** | **Vol 3 year** | **Vol 3 year** | **Vol 3 year** |
|---|---|---|---|---|
| **Excess Return on the Market** | | -0.0256** | -0.190** | -0.167* |
| | | (0.027) | (0.028) | (0.029) |
| **Small-Minus-Big Return** | | | -0.283** | -0.184** |
| | | | (0.039) | (0.040) |
| **High-Minus-Low Return** | | | 0.118** | 0.077 |
| | | | (0.041) | (0.053) |
| **Robust Minus Weak Return** | | | | 0.389** |
| | | | | (0.057) |
| **Conservative Minus Aggressive Return** | | | | 0.022 |
| | | | | (0.083) |
| **Momentum** | | | | -0.005 |
| | | | | (0.027) |
| **Intercept** | -0.004** | -0.003* | -0.003** | -0.004** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| **Number of Observations** | 597 | 597 | 597 | 597 |

*Note:* ** $p < .01$, * $p < .05$

month-long return measure is justifiable starting and ending in the same month. In the case of volatility however the weekly return measure better represents a month holding period and the returns generated. Regardless of this, Literature is split on the existence and Viability of the volatility anomaly and these mixed results fit into that body of work quite well.

### 6.4 Component Analysis of Z Score Returns

Being a composite ratio, the Z score represents more than a single measure, as a result, each of the components can have differing effects on the result obtained in the previous section. We can figure out in what capacity each of these components affects the return profile. This could give us a better understanding of why the return of Z score is hump shaped and whether any one of the ratios is a better bet than the Z score.

To achieve this result, we can regress the returns found by taking the difference between deciles, against similar return schedules for each of the 5 component ratios. The result for the same are included in Table 6. It shows what parts of the Z score return profile are influenced by each of its components. As expected, X3 (EBIT/Total Assets) is negative but X2 (Retained Earnings/Total Assets) is positive. Regardless both are insignificant which is consistent with why Z score was unexplained by the CMA factor. Their insignificance can also be chalked up to their movement in the period being in opposite directions. More importantly, it is perhaps the idea that these two ratios no longer express truly what they were initially included for. Being markers for growth and profitability, they were initially included in Altman's Z score due to their prevalence as important ratios in their respective categories. However changing business models may mean that firms

Table 7: **Z Score**

| | **Z Score** | **Z Score** | **Z Score** | **Z Score** |
|---|---|---|---|---|
| **Excess Return on the Market** | | -0.003<br>(0.045) | -0.009<br>(0.048) | 0.007<br>(0.051) |
| **Small-Minus-Big Return** | | | 0.087<br>(0.071) | 0.161$^{*}$<br>(0.075) |
| **High-Minus-Low Return** | | | 0.071<br>(0.071) | 0.001<br>(0.098) |
| **Robust Minus Weak Return** | | | | 0.302$^{*}$<br>(0.100) |
| **Conservative Minus Aggressive Return** | | | | 0.078<br>(0.151) |
| **Momentum** | | | | -0.041<br>(0.049) |
| **Intercept** | 0.020$^{**}$<br>(0.002) | 0.020$^{**}$<br>(0.002) | 0.020$^{**}$<br>(0.002) | 0.019$^{**}$<br>(0.002) |
| **Number of observations** | 596 | 596 | 596 | 596 |

*Note:* ** $p < .01$, * $p < .05$

nor investors value these metrics as much as they used to. Gavin (2011) puts more light on the issue writing that the ratios make firms seem more profitable than they are. Also interestingly, even though the X4 weightage in the Z score is so low at .006, the beta is nearly half of the dominant ratio X5 at .359. If looked at from the traditional notion of risk , this indicates that a higher buffer from equity to debt is riskier to a company. This also implies higher sales pose a greater risk in the case of X5. This is where the work of Tetlock and Schroeter (2010) is important. The authors posit that returns can also be derived from an undervaluation of fundamental strength. This argument was also partially accepted by Fama and French when they included the CMA factor in their work. The differing rankings as obtained by these sorts and the original equation indicates that a new Z score made of weights based on risk loadings of each of the factors might create a better factor for returns.

### 6.5 Liquidity Sorts

Now that we understand that these anomalies do indeed exist, before checking for their form and checking variability over time, we first need to verify if they truly exist and are not simply a result of the limits to arbitrage. Cementing whether they are independent of liquidity constrains can help us better link them to the underlying factor being used. Thus, dividing the stocks into liquid and non-liquid might help us verify if that is the case. In the case of Z score sorts, Table 10 shows that the anomaly persists through both

Table 8: **Z Score with Weekly Returns**

| | **Z Score** | **Z Score** | **Z Score** | **Z Score** |
|---|---|---|---|---|
| **Excess Return on the Market** | | -0.053** | -0.058** | -0.035* |
| | | (0.016) | (0.017) | (0.017) |
| **Small-Minus-Big Return** | | | 0.016 | 0.074** |
| | | | (0.025) | (0.025) |
| **High-Minus-Low Return** | | | -0.014 | -0.024 |
| | | | (0.025) | (0.033) |
| **Robust Minus Weak Return** | | | | 0.1226** |
| | | | | (0.034) |
| **Conservative Minus Aggressive Return** | | | | 0.017 |
| | | | | (0.018) |
| **Momentum** | | | | 0.035** |
| | | | | (0.017) |
| **Intercept** | 0.006** | 0.006** | 0.006** | 0.005** |
| | (0.001) | (0.001) | (0.001) | (0.001) |
| **Number of Observations** | 597 | 597 | 597 | 597 |

*Note:* ** $p < .01$, * $p < .05$

liquid and illiquid stocks. The results on Table 11 are disappointing as the anomaly exists only in illiquid stocks. However, it is to be noted that for liquid stocks, the intercept is negative pointing towards the low volatility anomaly. These results are insignificant, however. The opposite intercept values also show why the anomaly is so weak throughout the sample. The Z score returns being significant post liquidity can help us eliminate Dichev's claim that this was a case of illiquidity and while it does not prove that causality flows from distress to size, it's hard to believe the opposite is true.

### 6.6 Testing for Variability Over Time

There are two main ways to test for the effect of an event on anomaly returns. The first is to test for the returns themselves against an event. This method is the simpler one, however as De Groot (2017) points out, the average of the factor could respond evenly and thus the spread between the top and bottom deciles is not very likely to move.

#### *6.6.1 Portfolio sort regressions*

The results for the first of these methods are included in Table 12. The results are obtained using the regular RET measure for Z score and weekly returns for the volatilities. The results are significant only for the volatilities. Notably, the market downturns do in fact increase the low volatility anomaly as the beta is negative. When there is a contraction there is indeed a greater return in lower volatility stocks, but this result is insignificant. If significant, this would be in line with the work of Blitz and Van Vliet (2019) who found that the low volatility anomaly is more robust in downturns.

Table 9: **Returns Regressed Against Component Ratios**

| | **Z** |
|---|---|
| **X1** | $0.097^{**}$ (0.033) |
| **X2** | 0.071 (0.025) |
| **X3** | -0.040 (0.041) |
| **X4** | $0.251^{**}$ (0.021) |
| **X5** | $0.867^{**}$ (0.036) |
| **Intercept** | $0.002^{**}$ (0.001) |
| **Number of observations** | 596 |

*Note:* ** $p < .01$, * $p < .05$

Table 10: **Z Score Liquidity Sorts**

| | **Illiquid** | **Liquid** |
|---|---|---|
| **Intercept** | $0.036^{**}$ (0.003) | |
| **Intercept** | | $0.032^{**}$ (0.002) |
| **Number of observations** | 456 | 453 |

*Note:* ** $p < .01$, * $p < .05$

The opposite is true for the global financial crisis. The beta is positive and significant meaning that after the global financial crisis firms with higher volatility earned more. Now once again this could be due to the spread itself declining or a difference in risk appetites which led to firms with lower volatilities being priced higher which reduced the anomaly.

### *6.6.2 Fama Macbeth Regressions*

The next method to evaluate the variance of anomalies over events is Fama Macbeth regressions. These regressions allow us to see how the effect of a factor on return changes over time.

Table 11: **Volatility Liquidity Sorts**

| | **Illiquid** | **Liquid** |
|---|---|---|
| **Intercept** | $0.009^{**}$ (0.002) | |
| **Intercept** | | -0.004 (0.002) |
| **Number of observations** | 452 | 449 |

*Note:* ** $p < .01$, * $p < .05$

Table 12: **Regressions of Decile Differences with Event Dummies**

| | **Z Score Diff** | **Vol 1 Year Diff** | **Vol 3 Year Diff** |
|---|---|---|---|
| **gfc** | 0.004 (0.005) | $0.008^{*}$ (0.003) | $0.006^{*}$ (0.003) |
| **contraction** | 0.003 (0.006) | -0.005 (0.004) | -0.006 (0.004) |
| **Intercept** | $0.018^{**}$ (0.003) | $-0.007^{**}$ (0.002) | $-0.005^{**}$ (0.002) |
| **Number of observations** | 596 | 596 | 596 |

*Note:* ** $p < .01$, * $p < .05$

**Z SCORE**

Starting with the Z score the results are presented in Table 13. The results of the below table are like the results in Table 12, with very little to no effect of the events on the effect Z score has on the Return Variable. The interesting thing, however, is the betas for Z2 in all cases. Z2 is the square of Z score and as seen below has a negative beta. This curve was at the centre of Kim's work in 2016 and helps reconcile the ideas of Campbell et.al(2008) and the Fama French paper. Starting from the peak of the curve as we move towards a higher Z score, investors are rewarded with lesser and lesser return for taking on lesser risk. Left of the peak investors are being punished for investing in what is essentially a failing company and this has more instances of the gamble not paying off.

Looking at the actual relationship between RET and Z score, transforming to non-log terms and then deriving the equation to reach a peak, the max Z score is 3.625 (Figure 1). This is very close to Kim's Z score maximum of 5. Most importantly however we can evaluate with this regression the two sides of the risk return debate. Z score essentially corroborates what Campbell et.al write, where the higher risk of distress reduces returns. This is also in line with Lochstoer and Tetlock. With Z2's beta being negative, we can see what point where we are playing it too safe, and returns are no longer justified by

Table 13: **Z Score FMB regressions**

| | **RET** | **RET** | **RET** |
|---|---|---|---|
| **Z score** | 0.029** (0.002) | 0.029** (0.002) | 0.015** (0.001) |
| **Z2** | -0.004** (0.001) | -0.004** (0.001) | |
| **cons** | 0.002 (0.003) | -0.000 (0.003) | 0.007** (0.002) |
| **gfc** | | 0.002 (0.001) | |
| **gfc Z** | | 0.001 (0.000) | |
| **contraction** | | | -0.001 (0.002) |
| **contraction Z** | | | 0.000 (0.000) |
| **N** | 1921322 | 1921322 | 1921322 |

*Note:* ** $p < .01$, * $p < .05$

fundamental strength. Thus, we can take Z score to be the same effect of Campbells distress indicators and Z2 the effect as the same as the DEF factor of Fama French. That is not the say they are the same, but they probably represent each side of the debate. Seeing if there is in fact a peak that is consistent, we can cut the sample in half and work towards the same result in both. The result for the first 25 years of the sample is a maximum at 27.896. The second half is the Exact same as Kim's work with the maximum being at 5.03 (Figure 12 a & b). The wide variation over the peak brings up a very interesting point as to whether this means that the Z score is indicative of bankruptcy risk. While this is still the best estimation that we posses via ratio analysis, the result for the real relationship between distress risk and returns is perhaps somewhere in between. These values however even though in our dataset, cannot provide us with usable portfolio sorts as not enough companies have Z scores this high to create diversified portfolios. As shown by the histogram below, 98% of the data is between -1 and 1.5

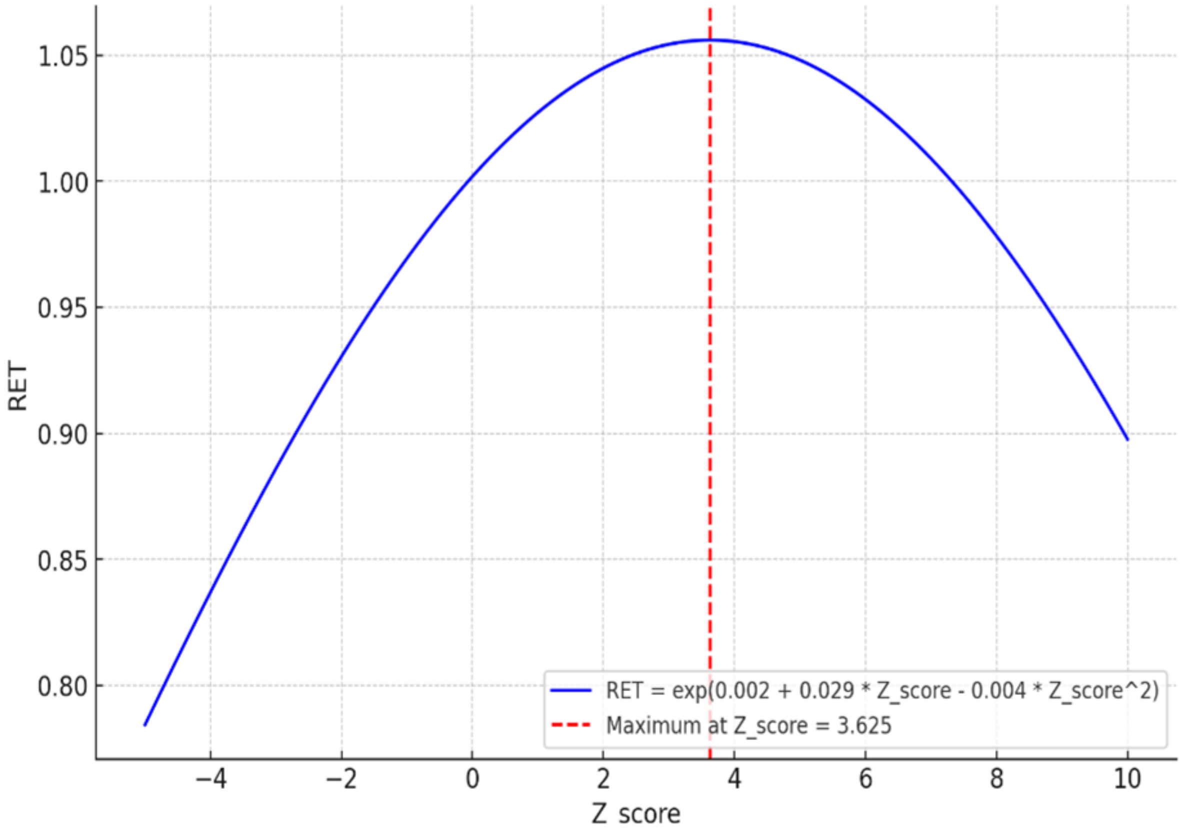



Figure 1: Z Score Graph

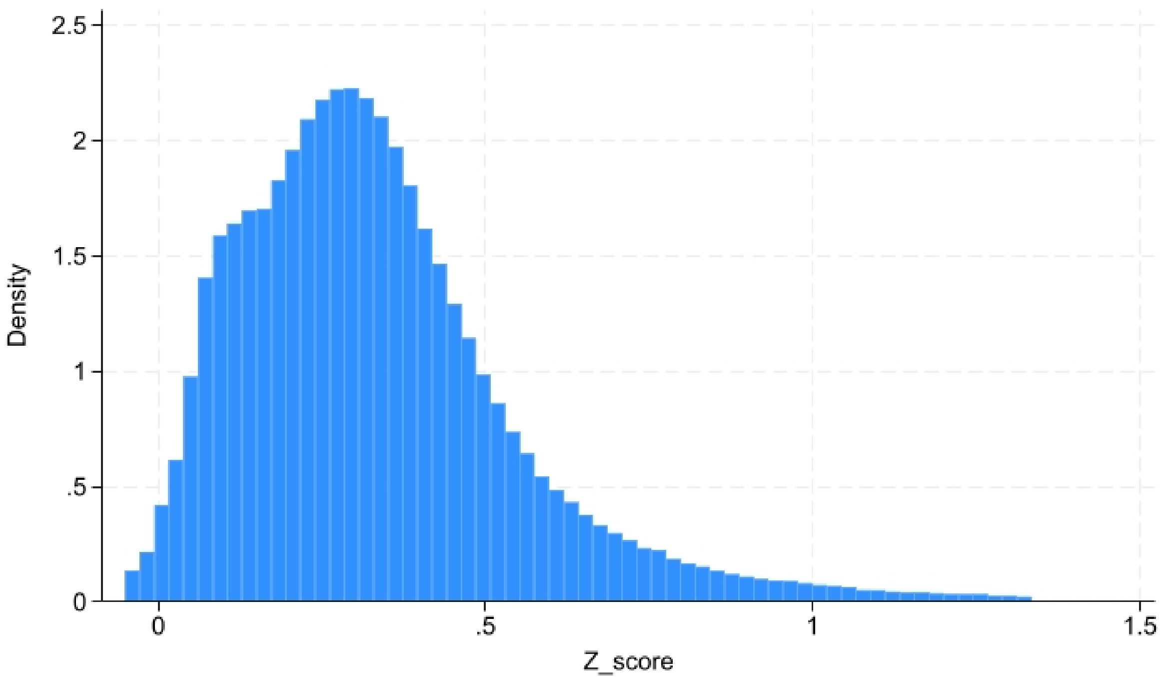



Figure 2: Z Score Histogram

**Volatility**

Table 14 shows us volatility against weekly returns. As we can see, both the interaction effects are insignificant. This along with the insignificance of the Interaction variables in Table 9 helps us disentangle what the reasoning behind the difference is. As it stands, the difference in returns before and after the Global financial crisis is due to the narrowing spread between the highest and lowest volatility deciles. The difference cannot be attributed to a difference perception or a higher "per unit of risk" return.

Table 14: **Volatility and Weekly Returns**

| | **Weekly Return** | **Weekly Return** |
|---|---|---|
| **Vol 1** | $0.032^{**}$ | |
| | (0.009) | |
| **GFC Vol 1** | 0.001 | |
| | (0.002) | |
| **Contraction Vol 3** | 0.002 | 0.002 |
| | (0.004) | (0.003) |
| **Cons** | -0.000 | -0.000 |
| | (0.001) | (0.001) |
| **Vol 3** | | $0.033^{**}$ |
| | | (0.009) |
| **GFC Vol 3** | | -0.001 |
| | | (0.002) |
| **N** | 1921322 | 1921322 |

*Note:* ** $p < .01$, * $p < .05$

### 6.7 Double sorts

The last part of this inquiry is to create deciles sorted based on both the Z score and Volatility of a stock to see if patterns emerge in their average returns. Table 15 displays the average returns for 25 of the 100 deciles calculated. A matrix consisting of deciles 1,3,5,7 & 10 of both factors.

Table 15: **Double Sorts with Volatility and Z Score**

| | **V1** | **V3** | **V5** | **V7** | **V10** |
|---|---|---|---|---|---|
| **Z1** | 0.00871 | 0.00411 | -0.00555 | -0.01857 | 0.00591 |
| **Z3** | 0.00851 | 0.00661 | 0.00394 | -0.00261 | 0.02449 |
| **Z5** | 0.00849 | 0.00897 | 0.00755 | 0.00731 | 0.03867 |
| **Z7** | 0.01008 | 0.01364 | 0.01232 | 0.01769 | 0.05035 |
| **Z10** | 0.01135 | 0.01785 | 0.02109 | 0.02563 | 0.07828 |

The patterns are quite similar looking at the corner portfolios, the revenue increases as volatility or Z score increases. There are however two things of note, for each level of Z score, the volatility rising first leads to lower returns before rising at the end in V10.

Secondly, the difference between highest to lowest Z score deciles increase as volatility increases. Hence, if there is indeed a positive relationship between return and Volatility, An ideal portfolio would be made of volatile companies that have the highest Z scores.

## 7 Conclusion and Discussion

The results of the above thesis have been fruitful in many regards. It has contributed to current literature in all 4 manners listed in the beginning. We now have tested for both anomalies on latest datasets, shown the form of the Z score anomaly & researched why it may exist this way. Volatility on the other hand was hard to verify but it did respond to the events studied lending credence to the idea that it is systematic. We can also definitively answer our main research questions:

1. Do the Z-score anomalies and Volatility anomalies still exist controlling for transaction costs?
The Z score anomaly has shown to exist beyond liquidity constraints for our sample. The volatility anomaly however has fallen short of being practical as it did not survive sorting by liquidity.
2. Do these anomalies show variance over time and over different phases of the business cycle?
The volatility anomaly does indeed change over time with the global financial crisis. The change is primarily due to a spike in average volatility which led to the spread between the high and low volatility which naturally led to the anomaly being stronger. The Z score anomaly does not waiver over the global financial crisis. Both anomalies were found to be non-respondent to the contraction and expansion of business cycles. It is to be said however that the periods of contraction were very low, and the sample was very unbalanced
3. Do the anomalies interact with one another?
The anomalies do indeed affect each other, more prominently, higher volatility leads to a higher return when sorting by the Z score of a stock.

We can also address the three criticisms of anomalies laid out at the start of this thesis. The first criticism stating that anomalies can be explained even by fundamental performance of a stock was indeed true to a point. While there was indeed a benefit to investing in a stock with great fundamentals, it is also important to look at the amount of risk being taken as an optimal portfolio would be a mix of both. The second argument which states that anomalies are only due to market inefficiencies is also true to a point as the volatility anomaly as we saw exists only in the illiquid stocks in the sample. The last argument which posits p-hacking is based on a study-to-study basis and checking for a larger time period and with different return measures I hoped to reduce this.

Besides these three questions, the shape of the Z score anomaly provided us a reconciliation between the two approaches to explaining anomalies. Finding the exact same peak at a Z score of 5 was indeed good, but a changing peak between different time periods makes it hard to say whether there is indeed a single peak. Differing business models and conditions can lead to differing ideas of what distress looks like and it might be hard to predict this peak outside of an observable sample.

This study brought up to me various further inquiries that could be made. The first relates to the Covid-19 pandemic. As shown above, a lot of the variables studied were shaken by the pandemic and had vast differences to pre covid levels. This could provide another event to be included into these regressions once enough time has passed for a balanced panel before and after. The next potential study is to use the returns of the DEF factor, which is a good proxy for default risk, to create a new Z score. By regressing the returns of the DEF factor against the returns of the component returns, it might provide us with the best weightages for each of the factors creating a Z score made from real risk of default.

## 8 References

Abraham, F., Cortina Lorente, J. J., & Schmukler, S. (2020). Growth of Global Corporate Debt: Main Facts and Policy Challenges. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3690997

Akcigit, U., & Ates, S. (2019). Ten Facts on Declining Business Dynamism and Lessons From Endogenous Growth Theory. SSRN Electronic Journal: https://doi.org/10.2139/ssrn.3467983

Al-Manaseer, Sufian, and Suleiman Al-Oshaibat. "Validity of Altman Z-Score Model to Predict Financial Failure: Evidence from Jordan." *International Journal of Economics and Finance*, vol. 10, no. 8, 24 July 2018, p. 181, https://doi.org/10.5539/ijef.v10n8p181.

Almaskati, Nawaf, et al. "A Horse Race of Models and Estimation Methods for Predicting Bankruptcy."
*Advances in Accounting*, vol. 52, Mar. 2021, p. 100513, https://doi.org/10.1016/j.adiac.2021.100513. Accessed 17 Nov. 2021.

Altman, E. I. (2020). FINANCIAL RATIOS, DISCRIMINANT ANALYSIS AND THE PREDICTION OF CORPORATE BANKRUPTCY. The Journal of Finance, 23(4), 589–609. https://doi.org/10.1111/j.1540-6261.1968.tb00843.x

Altman, Edward I., et al. "Distressed Firm and Bankruptcy Prediction in an International Context: A Review and Empirical Analysis of Altman's Z-Score Model." *SSRN Electronic Journal*, 2014, https://doi.org/10.2139/ssrn.2536340.

Anton, S. G., & Afloarei Nucu, A. E. (2020). The Impact of Working Capital Management on Firm Profitability: Empirical Evidence from the Polish Listed Firms. Journal of Risk and Financial Management, 14(1), 9. https://doi.org/10.3390/jrfm14010009

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. https://doi.org/10.1111/j.1540-6261.1977.tb01979.x

Baker, M. P., & Wurgler, J. (2015). The Risk Anomaly Tradeoff of Leverage. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2577948

Banz, R. W. (1981). The relationship between return and market value of common stocks. Journal of Financial Economics, 9(1), 3–18. https://www.sciencedirect.com/science/article/pii/0304405X81900180

Bellucci, A., Borisov, A., Gucciardi, G., & Zazzaro, A. (2022). The reallocation effects of COVID-19: Evidence from venture capital investments around the world. Journal of Banking &Finance, 106443. https://doi.org/10.1016/j.jbankfin.2022.106443

Bharath, S. T., & Shumway, T. (2008). Forecasting Default with the Merton Distance to Default Model. Review of Financial Studies, 21(3), 1339–1369. https://doi.org/10.1093/rfs/hhn044

Birge, J. R., & Zhang, Y. (2017). Default Risk Premia and a Non-Linear Asset Pricing Model. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.3061712

Blitz, D., & Vliet, van. (2019). The Volatility Effect: Lower Risk Without Lower Return. Journal of Portfolio Management, pp. 102-113 https://papers.ssrn.com/sol3/papers.cfm?abstract id=980865

Cao, M., & Wei, J. (2005). Stock market returns: A note on temperature anomaly. Journal of Banking & Finance, 29(6), 1559–1573. https://doi.org/10.1016/j.jbankfin.2004.06.028

De Groot, W. (2017). Assessing Asset Pricing Anomalies [Review of Assessing Asset Pricing Anomalies]. ERIM Electronic Series Portal; ERIM PhD Series in Research in Management. https://repub.eur.nl/pub/103490/EPS2017437F A9789058925015.pdf

Dichev, I. D. (1998). Is the Risk of Bankruptcy a Systematic Risk? The Journal of Finance, 53(3), 1131–1147. https://doi.org/10.1111/0022-1082.00046

Erhard, R. (2020). Explaining the Profitability Anomaly. https://doi.org/10.26226/morressier.5f0c7d3058e581e69b05d150

Ericsson, J., Huang, X., & Mazzotta, S. (2016). Leverage and asymmetric volatility: The firm level evidence. SSRN Electronic Journal, 38(A), 21. https://doi.org/10.2139/ssrn.964447

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. https://doi.org/10.1016/0304-405X(93)90023-5

Haskins, Mark E., and Phillip E. Pfeifer. "Predicting Corporate Bankruptcy." *SSRN Electronic Journal*, 2009, https://doi.org/10.2139/ssrn.1422923.

Horak, Jakub, et al. "Support Vector Machine Methods and Artificial Neural Networks Used for the Development of Bankruptcy Prediction Models and Their Comparison." *Journal of Risk and Financial Management*, vol. 13, no. 3, 24 Mar. 2020, p. 60, https://doi.org/10.3390/jrfm13030060.

Hou, K., Xue, C., & Zhang, L. (2018). Replicating Anomalies. The Review of Financial Studies, 33(5), 2019–2133. https://doi.org/10.1093/rfs/hhy131

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.

Jensen, M. C. (1978). Some anomalous evidence regarding market efficiency. Journal of Financial Economics, 6(2-3), 95–101.https://doi.org/10.1016/0304-405x(78)90025-9

Jenkins, D. S. (2003). The Transitory Nature of Negative Earnings and the Implications for Earnings Prediction and Stock Valuation. Review of Quantitative Finance and Accounting, 21(4), 379–404. https://doi.org/10.1023/b:requ.0000004784.30694.0c

Keim, D. B. (1983). Size-related anomalies and stock return seasonality. Journal of Financial Economics, 12(1), 13–32. https://doi.org/10.1016/0304-405x(83)90025-9

Kim, D. (2016). Size Premium , Distress Risk and Distress Anomaly. Semantic Scholar. https://www.semanticscholar.org/paper/Size-Premium-%2C-Distress-Risk-and-Distress-Anomaly-Kim/85e92e60837a7689360a0bdea951c8655384ff98

Koh, K. (Rachel), & Nawalkha, S. K. (2020). Firm efficiency and the investment anomalies. Managerial Finance, 46(12), 1589–1603. https://doi.org/10.1108/mf-07-2019-0380

Korol, Tomasz. "Dynamic Bankruptcy Prediction Models for European Enterprises." *Journal of Risk and Financial Management*, vol. 12, no. 4, 9 Dec. 2019, p. 185, https://doi.org/10.3390/jrfm12040185.

Leković, M. (2021). Historical development of portfolio theory. Tehnika, 76(2), 220–227. https://doi.org/10.5937/tehnika2102220l

Litzenberger, R. H., & Ramaswamy, K. (1979). The effect of personal taxes and dividends on capital asset prices. Journal of Financial Economics, 7(2), 163–195. https://doi.org/10.1016/0304-405x(79)90012-6

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

Liu, Y. (2018, September 27). Liquidity. Business Economics. Retrieved from http://hdl.handlenet/2105/45092

LOCHSTOER, L. A., & TETLOCK, P. C. (2020). What Drives Anomaly Returns? The Journal of Finance, 75(3), 1417–1455. https://doi.org/10.1111/jofi.12876

Mulder, E. (2020). The Low Volatility Anomaly Put to The Test. In Erasmus University Thesis Repository.

Ohlson, James A. "Financial Ratios and the Probabilistic Prediction of Bankruptcy." *Journal of Accounting Research*, vol. 18, no. 1, 1980, pp. 109–131, https://doi.org/10.2307/2490395.

O'donovan, J. (n.d.). Understanding the Asset Growth Anomaly. https://www.ivey.uwo.ca/media/3797068/understanding-the-asset-growth-anomaly.pdf

Pisula, Tomasz. "An Ensemble Classifier-Based Scoring Model for Predicting Bankruptcy of Polish Companies in the Podkarpackie Voivodeship." *Journal of Risk and Financial Management*, vol. 13, no. 2, 19 Feb. 2020, p. 37, https://doi.org/10.3390/jrfm13020037. Accessed 26 Apr. 2021.

Roy, A. D. (1952). Safety First and the Holding of Assets. In Google Books. University of Cambridge, Department of Applied Economics.https://books.google.ca/books/about/Safety First and the Holding of Assets.html?id=H87fHAAACAAJ &redir esc=y

Sharpe, W. F. (1977, January 1). THE CAPITAL ASSET PRICING MODEL: A "Multi-Beta" Interpretation (H. Levy & M. Sarnat, Eds.). ScienceDirect; Academic Press. https://www.sciencedirect.com/science/article/abs/pii/B9780124458505500118

Treynor, J. L. (1961, August 8). Market Value, Time, and Risk. Papers.ssrn.com. https://papers.ssrn.com/sol3/papers.cfm?abstract id=2600356

Vedy, & Santoso, H. (n.d.). Asset Turnover, Profitability, Capital Structure, and Firm Value in Coal Subsector in Indonesia. Economics and Accounting (TiCMEA), 1, 2022. https://feb.untan.ac.id/wp-content/uploads/2023/02/39-1.pdf

Vuolteenaho, Tuomo. "What Drives Firm-Level Stock Returns?" *The Journal of Finance*, vol. 57, no. 1, Feb. 2002, pp. 233–264, https://doi.org/10.1111/1540-6261.00421.

Wang, B. (2017). The Cash Conversion Cycle Spread.
SSRN Electronic Journal. https://doi.org/10.2139/ssrn.2964330

# A Appendix

## A.1 Time Series Graphs

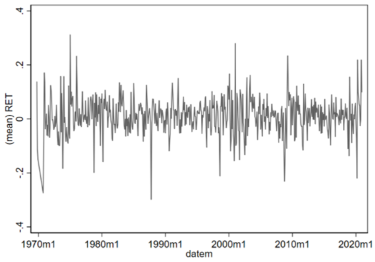



Figure 3: Graph 1.0 - Returns

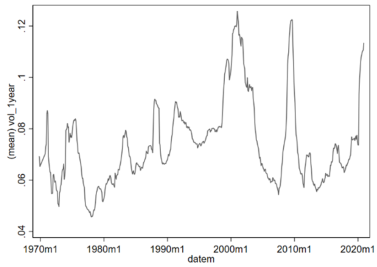



Figure 4: Graph 1.1 - Volatility

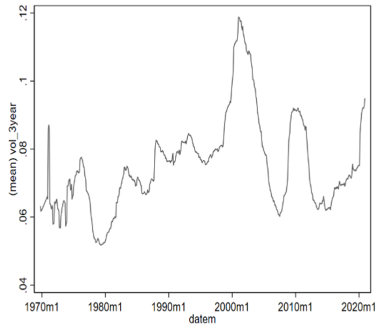



Figure 5: Graph 1.2 - Volatility 3 year

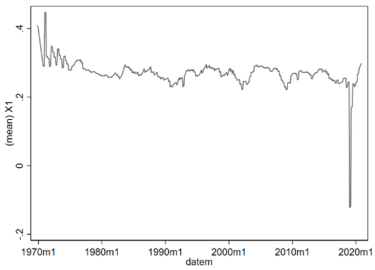



Figure 6: Graph 1.3 - X1

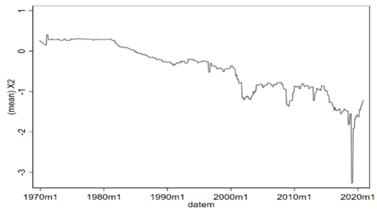



Figure 7: Graph 1.4 - X2

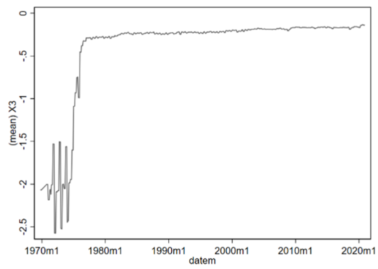



Figure 8: Graph 1.5 - X3

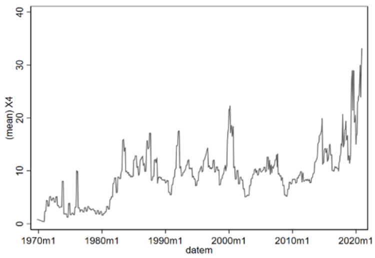



Figure 9: Graph 1.6 - X4

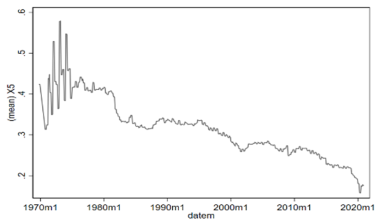



Figure 10: Graph 1.7 - X5

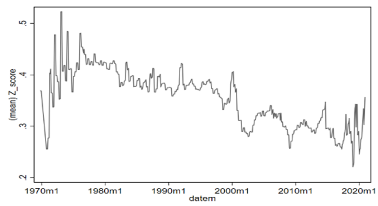


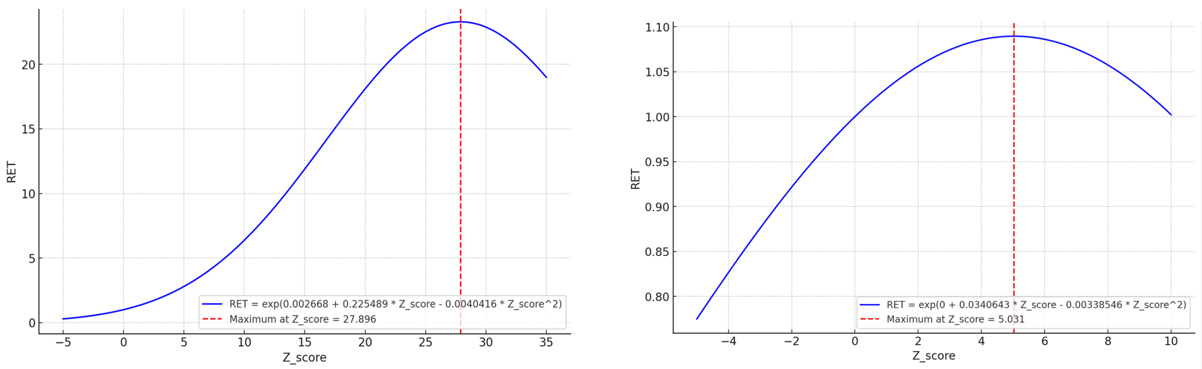

Figure 11: Graph 1.8 - Z score

## A.2 Peak Return Profiles Before and After 1995

Table 16: **Z Score Returns Post 2000**

| | Z Score | Z Score | Z Score | Z Score |
|---|---|---|---|---|
| **Excess Return on the Market** | | -0.113 | -0.148 | -0.109 |
| | | (0.084) | (0.088) | (0.104) |
| **Small-Minus-Big Return** | | | 0.171 | 0.222 |
| | | | (0.129) | (0.149) |
| **High-Minus-Low Return** | | | 0.108 | -0.011 |
| | | | (0.118) | (0.172) |
| **Robust Minus Weak Return** | | | | 0.154 |
| | | | | (0.184) |
| **Conservative Minus Aggressive Return** | | | | 0.163 |
| | | | | (0.253) |
| **Momentum** | | | | -0.017 |
| | | | | (0.083) |
| **Intercept** | 0.028 | 0.029 | 0.028 | 0.027 |
| | (0.004) | (0.004) | (0.004) | (0.004) |
| **Number of Observations** | 251 | 251 | 251 | 251 |

*Note:* ** $p < .01$, * $p < .05$




(a) Peak Returns Before 1995

(b) Peak Returns After 1995

Figure 12: Comparison of Peak Returns Before and After 1995