26 JULI 2021

BUFFETT'S INDICATOR AND SHILLER'S CAPE RATIO AS SENTIMENT PROXIES

IN SEARCH OF INVESTOR SENTIMENT STRATEGIES

Michael Muit (469889)

Erasmus School of Economics – Master Financial Economics

Supervisor : Dr. Laurens Swinkels

Second assessor : Dr. Roy Kouwenberg

Abstract

This paper attempts to show the predictive power of the Buffett indicator and Shiller CAPE ratio. This paper first investigates the ability of the Buffett indicator and CAPE ratio to be sentiment proxies and finds that it can be a good proxy for sentiment in a similar fashion to the VIX but not compared to a more sophisticated index such as the one by Baker and Wurgler (2006). Next this paper seeks to test the predictive power on six well documented anomalies which comprise the six-factor model across different countries in and outside of the USA. This paper finds that certain anomalies in certain countries do have a positive relation with the Buffett indicator and the Shiller CAPE ratio. However, when looking at the complete picture one cannot conclude there to be a significant positive relation with all anomaly returns. A trading strategy using Shiller's CAPE ratio or Buffett indicator does not show any real promise either for most of the anomalies , with a standout Sharpe ratio for the trading strategy which uses the Buffett indicator to predict the Cash-based profitability anomaly, with a value of 0.629 compared to the market Sharpe ratio of 0.322. Overall, this paper adds to the scarce literature on the predictive power of the Buffett indicator and is the first to take a further look into its and the Shiller CAPE ratio's ability to predict anomaly returns. This paper does not show many significant results, but still provides a good starting point for further research into the topic.

1) Introduction

In recent years research into and interest in mispricing, and investing strategies that use this mispricing, has substantially risen. Market prices currently seem very high, and overpricing seems to exist, some articles even talk about a new bubble being created (Martin, 2021). With a shift in interest trying to explain mispricing and mostly overpricing, Baker and Wurgler (2006) tried creating a proxy to capture investor sentiment and found evidence that indicate high investor sentiment can be exploited. They suggest that classical financial theories should incorporate a sentiment proxy to be able to better predict prices via pricing models. Ever since Keynes (1936), many have tried to find a market-wide accepted sentiment proxy that causes prices to depart from fundamental values, but literature has never found a single, widely accepted answer to the question, which investor sentiment proxy is the best.

Another way in which investors try to profit from the market is via anomalies. One of the most studied and in practice used anomalies is momentum. Momentum strategies consist of a relation between stocks near future returns and relative recent returns history. Suggesting stocks that performed well in the near past will perform well in the following period and vice versa. Many traditional finance researchers have tried to find a logical explanation, but no traditional asset pricing models such as Fama and French's (1993) three factor model or the CAPM could explain the phenomenon.

In the last decade, a few researchers have tried to find links between such anomalies as momentum and investor sentiment. Stambaugh, Yu and Yuan (2012) used Baker and Wurgler's (2006) sentiment index to test the effect of the index on 11 anomalies, including momentum, and found a noticeable effect on returns, mostly due to overpricing in months following high sentiment. Other anomalies researched in this paper were, for example the profitability and investment anomalies, which later went on to become integral parts of asset pricing model. Two more interesting anomalies that have been a part of all major asset pricing models in recent years are the size and value anomaly, both of which are not included in the paper by Stambaugh et al (2012), but seem interesting to explore due to the huge predictive powers they hold. These 4 anomalies also find no theoretical explanation in the CAPM and make up the most important parts of the Fama and French five factor model. Therefore, just like momentum, it is interesting to see the effect of a sentiment proxy on these anomalies. Furthermore, most sentiment indexes, like the Baker and Wurgler index, are made up of US investor sentiment and focus mainly on variables on a smaller scale. This is where a macro sentiment proxy, which holds similar properties across different countries, can be an interesting addition as well as an extension to other markets outside the US.

Warren Buffett is known as one of the most successful investors in the world. He and his company, Berkshire Hathaway, have yielded great results in a period spanning multiple decades, even in the tumult of last year, it seemed like he and Berkshire Hathaway profited (Henderson, Financial Times 2020). Buffet's views on the stock market are therefore respected in the world of finance and a lot of his knowledge is shared and researched. In 1999 and 2001 Buffett was interviewed on his views on the stock market and presented a new indicator that could help explain stock returns and mispricing in the last few decades before his two interviews. This indicator was defined as the ratio of the market value of equity to economic output, in this case specifically GDP (Gross Domestic Product). He said that this indicator may be "the best single measure of where valuations stand at any given moment". Therefore, this indicator can be seen as another proxy for investor sentiment and should be a good sign for when the market is overpriced.

Another interesting ratio that has gained a lot of interest in the last 20 years, has been Shiller's P/E ratio or CAPE ratio (Cyclically Adjusted Price to Earnings ratio). Just like the Buffett indicator, the CAPE ratio can indicate market overpricing (Shiller,1996). Over the last 20 years, the CAPE ratio has been used by many investors to assess the state of the market, as a measure to see whether the market is over- or underpriced at a given point in time. However, much research into the predictive power of the CAPE ratio or its value as a sentiment proxy has not been conducted.

In this paper for a period from 1971 up until 2019, both the CAPE and the Buffett indicator of various countries will undergo various tests. First a test comparing the US Buffett indicator and the US CAPE ratio to well-established sentiment proxies will be conducted. After which for both the Buffett indicator and the CAPE ratio across various countries, the US, UK, Netherlands, France, Germany, Canada and Japan, a trading strategy will be constructed and tested. This trading strategy will focus on six anomalies, which comprise the Fama and French six factor model (2016): the size, value, momentum, cash-based and non-cash-based profitability, and investment anomaly. Finally, regressions will be run to test the predictive powers of the Buffett indicator and CAPE for the same countries on the same anomalies.

All this leads to the research question being:

"Is the Buffett indicator and/or the Shiller CAPE ratio a good proxy for investor sentiment and can it predict returns for five of the most well documented anomalies (size, value, momentum, profitability and investment factors) for the US, UK, Netherlands, France, Germany, Canada and Japan? "

This paper will focus on expanding the use of the Buffett indicator and the CAPE ratio. The paper tries to improve on the findings of Umlauft(2020), who provided the first evidence of predictive power of the Buffett indicator and the indicator's ability to evaluate whether the market is over- or underpriced. However, the paper by Umlauft (2020) does not compare its results to any other sentiment indices, nor does it show its predictive power for anomalies, nor does it provide any interesting trading strategies using the Buffett indicator and anomalies. All these previously unhandled points are handled in this paper, which is what makes this paper stand out. Furthermore, this paper elaborates on the findings of Stambaugh, Yu and Yuan (2012), who found that due to being able to successfully identify when a market is over- or underpriced, that the sentiment index by Baker and Wurgler (2006) could help predict anomaly returns. Following these findings and due to the natural ability of both Shiller's CAPE ratio as well as the Buffett indicator, to evaluate the state of the market, this paper will extend on the paper by Stambaugh, Yu and Yuan on whether any market evaluation tool, such as the Shiller CAPE ratio and the Buffett indicator, can help predict anomaly returns.

The rest of the paper will first focus on the theoretical framework, where the basis of investor sentiment's influence on anomalies and predictive power of the Buffett indicator and CAPE ratio will be presented. After which data and methodology will be handled, where the sample creation and various tests will further be explained. Subsequently the results following these tests will be discussed. At the end, a conclusion will follow, where limitations and further implications will also be discussed.

2) Theoretical framework:

2.1) Mis- and overpricing

In the world of finance, most work is written aiming to understand the stock market better and in which ways one can exploit any inefficiencies. Fama (1970) developed the efficient market hypothesis, meaning prices should reflect all information available and investors will be rational in order to maximize their wealth and utility. Furthermore Lintner (1965) and Sharpe (1964) assume with their theoretic model, the CAPM (Capital Asset Pricing Model), that all investors have similar outlooks on future returns and earnings even under uncertainty. Ever since the efficient market hypothesis and the CAPM have come to live, a lot of critical papers have been written in response.

Miller (1977) shows it is possible for prices to siege past fundamental values. He also concludes that, just like mentioned before, prices in normal circumstances will go down due to experienced investors shorting it or taking advantage of this mispricing, however limits to arbitrage, like short selling restraints, make this harder. He further shows, due to the uncertainty in the market, that it is impossible for all investors to have similar outlooks on the future, since all people have different interpretations of the same information. More papers on the effect of limits to arbitrage have been written since then, such as the papers by De Long et al. (1990) and Vishny and Schleifer (1997). In the first of the above-mentioned papers, they find risk, created by inexperienced investors, to deter arbitrageurs of betting against them, which leads to overpricing. Noise trading can thus lead to overpricing, due to its massive uncertainty. In that paper the fundamentals for the effect of sentiment on financial theories is found for the first time. Vishny and Shleifer show a more real-world view on how arbitrageurs are limited in correcting the market. They conclude that arbitrageurs are mostly investors with outside capital and the risk of needing to liquidate positions can, in very volatile situations, deter arbitrageurs from taking positions of arbitrage. Both these papers show how the working of the market itself can cause mispricing, which will mostly be overpricing according to Miller (1977). More recent papers further acknowledge the existence of mispricing and then mostly overpricing. Chu, Hirshleifer and Ma (2020) show, by using a database that tested relaxed short-selling constraints, that mispricing exists. They even link such mispricing to returns of 11 well-known asset pricing anomalies.

Another way in which mispricing is noticed, is due to behavioural effects. Tversky and Kahneman (1994) show that uncertainty can influence behaviour drastically and demonstrates people are susceptible to be influenced by certain heuristics and biases, which in turn causes irrational investment decisions. Kahneman and Riepe (1998) further write about his earlier findings with Tversky and how to prevent falling for certain biases and heuristics. They showed that rationality in investors should not be completely disregarded, but also that irrational behaviour is more prominent and easier to fall for than most traditional finance and efficient market hypothesis partisans seem to believe. Thaler and de Bondt (1985) write one of the most cited articles when it comes to behavioural finance and investing, cited over 10 000 times. Their paper focuses on finding whether stock markets and prices can overreact and concludes this to be the case. Which shows that just like in Miller (1977), prices can be higher than fundamental values. In more recent years a few more papers have linked behavioural effects to overpricing. Daniel and Hirshleifer (2015) show that investors can disagree on prices and the importance of certain information due to overconfidence, which in turn can cause mispricing. When, for example, an investor interprets news as negligible, he might overvalue a stock and inherently cause overpricing. Barber and Odean (2008) further show that investors are more enticed to buy attention-grabbing stocks, that experience more news, and are therefore bought more often, which in turn can lead to mis- and overpricing. This all shows that just like in the paper by Miller (1977), prices can siege past fundamental values.

2.2) Investor Sentiment

In more recent years, papers on behavioural finance have been focusing on using sentiment to explain over- and underreaction. In Daniel, Hirshleifer and Subrahmanyam (1997) they develop a theory using investor overconfidence and confidence shift due to self-attribution to predict over- and underreaction. The self-attribution bias is a psychological concept, where people tend to claim their successes as personal skill or wisdom, whilst they blame external factors for their losses. As an example, they used the reception of news, where people tend to overreact to private information whilst underreacting to public information. Further extension on psychological influences on over- and underreaction came a year later in a paper by Barberis, Shleifer and Vishny (1998). They present a model of investor sentiment concerning investor expectations on future earnings. They found investors to overestimate certain news announcements by not taking statistical values into account. They find stock prices to overreact to consistent patterns of good or bad news, even though these patterns carry low statistical weights. Both these papers show that over- and underreaction and investor sentiment are linked.

Since the first papers written about investor sentiment, a few papers have surfaced trying to create the best sentiment proxy to use. One of the more commonly used sentiment indexes is the one by Baker and Wurgler (2006 and 2007). Where in the first of the two papers (2006) they focus more on the explanatory power and construction of the index, in the second paper (2007) they focus more on the general state of sentiment indexes and their motivation for the one used in the paper from 2006. They use a sentiment proxy consisting of six measures, which are closed-end fund discount, number and first-day returns of IPOs, NYSE turnover, equity share in total new issues and dividend premium.

In their paper from 2006 they find there to be significant evidence that this proxy influences the cross-section of future stock returns. In their paper from 2007, they further describe as to why these six measures were chosen. First of all, they mention that investor sentiment is not straightforward to measure, but there is no fundamental reason why one cannot find imperfect proxies. Most of the data used for their index is chosen due to availability of this data and has been adjusted to correct for any idiosyncratic risk and economic fundamentals. Furthermore, they conclude that all aspects of the index on their own would also display similar, if not better results as the sentiment indexes combined.

As we can see in the literature above a lot of research exists on individual stock mispricing and its link to behavioural investing and investor sentiment, mostly for US stocks. However, when trying to see it on a more macro scale, not a lot of financial literature is available on the effect of macro sentiment on indices. The only paper close to such research is the paper by Baker, Wurgler and Yuan (2012). In this paper they look at investor sentiment in the UK, France, Germany, Japan, Canada, and a global index. Through various regression tests they conclude that investor sentiment on a global scale is a contrarian predictor of market time series results, whilst local scale investor sentiment is both an economic and statistically significant predictor of market time-series results. They conclude that the effects in all these countries and the global indices, follow similar patterns as the U.S. sentiment index and are of similar influence for pricing models.

2.3) Buffett indicator and Shiller CAPE ratio

In two interviews in 1999 and 2001 Warren Buffett gives his two cents on the stock market and where it stood at that time. In the first interview he shows how the stock market had grown in a 17-year period starting in 1982. He concludes that there are three things that might cause stock profits to grow. Firstly, interest rates falling, since this will make stocks more attractive, and more people will then join the market. Secondly, the rise of corporate profits as a percentage of GDP (Gross Domestic Product), otherwise stock profits can only be as large as GDP profits. Third and finally investing in companies that have a clear competitive advantage and not necessarily investing in the best firm in a booming industry, since tides can turn more often than expected. In the second interview he focuses more on the explanatory power of his assumptions and its forecasting ability. He gives financial advice about when to buy and not to buy, he advises to sell for high values and buy at low values of the ratio of market capitalization to GDP. With this interview he sparked interest in such ratios being able to notice major market mispricing, which as mentioned earlier, seems to also be heavily linked with sentiment. This all leads to some financial literature looking at the so-called Buffett indicator or the ratio between market-value-of-equity-to-GDP. It must be noted that in general, any type of ratio between market values and national product can be seen as a form of a Buffett indicator, but the one mentioned above is the most used in financial research.

However, unlike financial literature on sentiment in general, there has been relatively little research into the explanatory power of the Buffett indicator. Chang and Pak (2018) summarise some findings in financial articles about the Buffett indicator and looking into the relation between the Buffett indicator and GDP for numerous countries. First of all, they find that in general it is considered a buy-signal if the indicator, compared to historical data, is down two standard deviations, whilst it is considered a sell signal when it is up two standard deviations. Furthermore, they find a superlinear scaling relationship between the Buffett indicator and the GDP in a group of 34 countries, where they find the strongest effect for the 13 lowest GDP countries.

Umlauft (2020) improves upon research into the Buffett indicator in a working paper and did so successfully. In his paper he focuses on the US market when it comes to the market value of equity, GDP, and stock returns, where he uses the S&P 500 (Standard and Poor) returns. He does all this in a long period from 1951 up until 2019 and used quarterly data. For this period and its data, he concludes that the ratio of market-value-of-equity-to-GDP is a good tool for evaluating whether the stock market is over- or underpriced at any given moment. He also concludes that this indicator is consequently a good investor sentiment proxy towards the stock market, but also incorporates fundamental aspects as well. Due to the conclusions of this paper, one may seem to expect similar trends as to other sentiment indexes such as the one by Baker and Wurgler (2006, 2007), since there are some similar proxies used in these indices that also influence the Buffett indicator value, mostly the market value side of the ratio. For example, the number of IPO's influences the market value, since there are now more companies that bring value to the market, which in turn influences the Buffett indicator.

Furthermore Shiller (1996) created one of the most well-known ratios for the assessment of returns, namely the CAPE ratio. The CAPE is usually calculated as follows: divide stock/market prices by earnings by the average of the last 10 years of earnings per share. Shiller (1996) further concludes that the equity market is not a random walk model, which can contradict the efficient market hypothesis for non-time-varying equity risk premiums and therefore created this ratio. In financial literature the Buffett indicator is usually also more on a long-term scale compared to other sentiment proxies, so a comparison between these two would work well, since the CAPE ratio is also usually considered as a long-term investment tool. Dang and Xu (2018) further investigate using the Shiller CAPE as a sentiment proxy and found confirming results if orthogonalizing the CAPE in a similar way as the Baker and Wurgler sentiment index. In this paper however the focus lies more on the un-orthogonalized CAPE ratio, due to the comparability with the Buffett indicator.

This all makes it seem like both the Shiller CAPE and Buffett indicator are good proxies for overpricing and therefore could show a similar positive trend to the Baker and Wurgler sentiment index, which leads to hypothesis 1.A being:

"Buffett's market-value-of-equity-to-GDP indicator and Shiller's CAPE ratio follow a similar trend to the sentiment index by Baker and Wurgler index from 2006 and 2007. "

2.4) Volatility

Another proxy to gauge sentiment which has been studied in research in the last 15 years are volatility indices. A lot of research focuses more on using VIX (Volatility index) to measure other helpful additions to sentiment proxies or use the VIX as a comparative tool. Research, such as for example Smales (2014), show a negative relation between sentiment proxies and the VIX, since for periods with low volatility, prices are usually above fundamental values, while sentiment proxies show overpricing for high values. Barroso and Santa-Clara (2015) further find that volatility is also linked to many anomaly strategies and the state of the market, just like the sentiment index by Baker and Wurgler, but is more of an indicator instead of an index, just like the Buffett indicator and Shiller's CAPE ratio. For example, they find that a momentum strategy yields better results when taking volatility into account. Which all leads to hypothesis 1.B being:

"The Buffett indicator and Shiller's CAPE ratio follow an opposite trend to volatility"

2.5) Anomalies

Furthermore, as mentioned earlier, the market does not always seem to be efficient, and possibilities of high excess returns, which do not solely find their returns in risk, exist. Another way in which stocks show excess returns which cannot be explained by traditional asset pricing models, are anomalies. As mentioned, two of the more known asset pricing models are the CAPM and three factor model, these can explain a lot of anomalies but not all.

In this paper focus will lie on five anomalies relative to the CAPM, which are the momentum, investing, size, value, and profitability anomaly. These are some of the most well-known and among the most studied anomalies, which is why this paper will focus on these anomalies.

First of all, Jegadeesh and Titman (1993) were the first to properly explore the possibility of a so-called momentum strategy. In a sample period from 1965 up until 1989, for the NYSE and AMEX stock markets, they test whether recent past returns, in a period of 3 to 12 months, hold any predictive power for returns in the following 3 to 12 months. They find that holding previous winners and selling past losers realises abnormal returns in this period for these specific markets. To further extend on their findings Jegadeesh and Titman repeated their paper in 2001 and 2011 to see the results of the momentum anomaly in the 90's and 00's and find similar results, again showing momentum being one of the strongest pieces of evidence against the efficient market hypothesis. In traditional pricing models, momentum is such a major outlier and so difficult to explain, that Fama and French (2018) created a new model. This new model is called the six-factor asset pricing model, where a momentum component has been added to further improve this model. This model is currently seen as one of the most accurate pricing models, but to date no real explanation exists on how momentum can generate such abnormal returns, and this causes much discussion in financial literature.

Another interesting anomaly where traditional finance has not found an answer yet, is the gross profit premium by Novy-Marx (2010). He finds profitable firms to outperform unprofitable firms in a similar fashion as value stocks outperforming growth stocks, found in the paper by Stattman (1980) and Rosenberg, Reid and Lanstein (1985). Just like for the momentum anomaly, the profit anomaly is now added to be part of the six-factor model to help improve its predictability and just like the momentum anomaly, in financial literature there currently is no real explanation yet to how the profit anomaly creates its abnormal returns.

Two more parts of the six factors, that have been around the longest and are to this day still an important part of asset pricing models, are the size and value anomalies. In their original three factor model, which at the time was seen as the best return predicting model, Fama and French (1993) already included size and value to help explain abnormal returns. They found a model including factors for both size and value to explain 90% of excess returns at that point in time. Since then, due to the public knowledge of these anomalies and diversification of portfolios, the predictive power of these two anomalies has decreased. However, they are still part of the six-factor model and will always hold some predictive power, therefore being interesting enough to research.

Lastly Titman, Wei and Xie (2004) observe another anomaly concerning investments, they conclude that there is a negative relation between abnormal capital investments and future stock returns. They further notice there to be no link between this effect and risk or firm characteristics. To further improve upon the investment anomaly Xing (2008) deduces that, just like the profitability anomaly, the investment factor contains similar information that can help explain the value effect found by Stattman (1980) and Rosenberg, Reid and Lansteing (1985). Just like the profitability anomaly, the size anomaly, the value anomaly and the momentum anomaly, the investment anomaly is now also added into the six-factor model by Fama and French as a new component and its abnormal returns are also yet to be explained.

Most of these anomalies and pricing models are tested in a US only based setting. To further look across other countries a paper by Hanauer (2020) offers the solution. In this paper he tests multiple global factor models and found the model by Barillas, Kan, Robotti and Shanken (2019) to offer the best results, which is an adaptation to the Fama and French (2018) six factor model. He further concludes that for this model to be the best, it absolutely needs a momentum component and a profitability component, which are both included in that paper as well. However, instead of the profit-to-assets as the profitability factor, Hanauer (2020) concludes that a cash-based profitability factor is better.

These anomalies provide an interesting perspective for future investment strategies, since exploiting these anomalies can lead to large abnormal returns. Investors therefore always try to use these anomalies in their investment strategies. However, since not much is known about the trends they follow and what affects these anomalies, investors might skew away too early from these anomaly strategies due to ambiguity and could therefore miss out on the upsides of strategies using these anomalies. This is where the next paper comes in and tries to help explain these phenomena. In the paper, Stambaugh, Yu and Yuan (2011), research the effect of investor sentiment on 11 anomalies, including three of the five mentioned above, namely the momentum, profitability, and investment. In this paper they use the Baker and Wurgler (2006) sentiment index as a proxy for investor sentiment. They tested it in a setting in a period from 1965 to 2008. They found there to be greater profitability after months with high sentiment for 8 out of the 11 anomalies in a benchmark adjusted model. Where momentum and profitability showed to be two of the most influenced anomalies by investor sentiment. The investment-to-asset anomaly however does not show strong evidence of being influenced by investor sentiment. They further find, using regressions, the Baker and Wurgler sentiment index to be positively related to anomaly returns, since again, when the investor sentiment is high and the market is overpriced, they find anomaly returns to flourish.

This is where the Buffett indicator could be a good addition to try and see its predictive power on anomalies, just like the sentiment index by Baker and Wurgler is. Hypothesis 2a therefore is:

"The Buffett indicator can predict anomaly (size, value, momentum, profitability and investment factors) returns across different countries (the US, the UK, the Netherlands, France, Germany, Canada and Japan). Therefore, following months with high investor sentiment, excess returns should be higher for these anomalies. "

And hypothesis 2b will be:

"The CAPE ratio can predict anomaly (size, value, momentum, profitability and investment factors) returns across different countries (the US, the UK, the Netherlands, France, Germany, Canada and Japan). Therefore, following months with high investor sentiment, excess returns should be higher for these anomalies. "

3) Data and Methodology:

3.1) Data Buffett indicator

Data used in this research has been collected from various sources and over various time points, which will be discussed in this section of the paper. This data has been edited and eventually been used to get the results discussed in the following section.

First, let me discuss the data collection process for the different Buffett indicators. For the following countries data has been collected: the US, the UK, the Netherlands, Canada, Germany, Japan, and France. Out of these countries the easiest to collect historical Buffett indicator data was for the US. The original Buffett indicator mentioned by Warren Buffett in his interviews was the ratio between the market cap of the Wilshire 5000 divided by the US GDP, this also the Buffett indicator used in this paper for the US. Historical data for this indicator has been obtained from longtermtrends.net, for a monthly period stretching from April 1971 up until December 2019. The summary statistics for this Buffett indicator can be seen in table 1 and shows similar results to the Buffett indicator of the UK with a mean of approximately 0.8 and shows similar minimum and maximum values compared to the other Buffett indicators in general.

For all the other Buffett indicators there was not a similar source, where historical data could be gathered. Therefore, these Buffett indicators had to be calculated from the raw data of which Buffett indicators are composed of, namely the market capitalization (market cap) of a particular country divided by its GDP. Data for the raw values of the monthly market caps in US dollars for the UK, the Netherlands, Canada, Germany, Japan, and France has been obtained from DataStream for the period of April 1971 up until December 2019, provided by the Erasmus Data Service Centre (EDSC). Unfortunately for the Netherlands, France, Germany, Canada, and Japan data was only available starting from January 1973. Yearly GDP data in US dollars for the UK, the Netherlands, Canada, Germany, Japan, and France has been obtained from The World Bank data base in a period from 1971 up until 2019. After this data had been obtained, it was used to calculate the monthly Buffett indicator values, as mentioned, by dividing the market cap values by the GDPs of the countries. The summary statistics for the UK, the Netherlands, Canada, Germany, Japan, and France can be seen in table1. In table 1 one can see that almost all Buffett indicators have a mean value between 0.5 and 0.8, with only France and Germany being slightly off. When it comes to maximum values, a similar pattern can be seen, where only France and Germany seem to not follow the highs that other countries experience, but when looking at the minimum values one can see that France and Germany also have lower minimum values than the others, except for Canada which also shows a very low minimum value for the Buffett indicator.

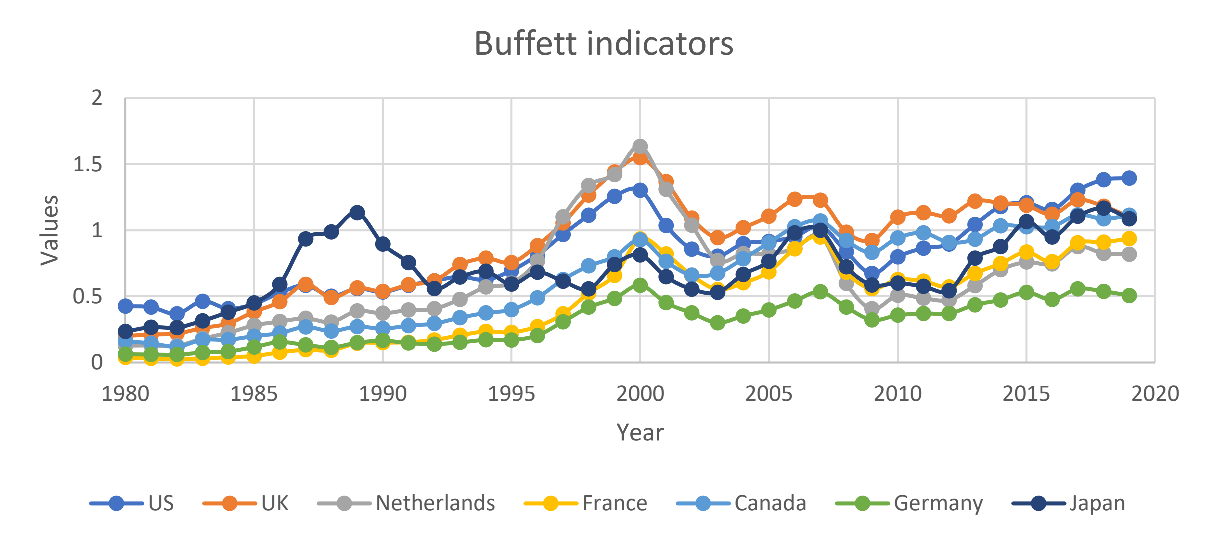
Furthermore, when looking at figure 1 one can observe that most of the Buffett indicators follow a similar pattern, with peaks around the dot.com bubble and the 2008 crisis. Keeping in mind the thoughts behind the Buffett indicator, the timing of these peaks seems logical, due to high values being associated with overpricing and low interest rates, which in turn is associated with bubbles which can burst such as in 2000 and 2008. For the main part, the graph in figure 1 shows similar features to table 1, where Germany shows the lowest curve, whilst the Netherlands shows the highest high around the year 2000, midst of the dot.com bubble.

However, when compared to the US Buffett indicator, one can see that most of the other countries have slightly different values and therefore the initial interest in exploring the predictive power of the Buffett indicator for countries outside of the US, looks to be of importance. Furthermore, it is of importance to note that the US Buffett indicator will not be orthogonalized in a similar way as the sentiment index in Baker and Wurgler (2007), due to comparability with the other countries, since the orthogonalization process cannot be performed for these other countries.

Table 1**:** Summary statistics of Buffett indicators for 7 countries in a period stretching from April 1971 up until December 2019.

Nation	**Observations**	**Mean**	**Standard deviation**	**Minimum**	**Maximum**
United States	585	0.774	0.303	0.333	1.488
United Kingdom	585	0.771	0.422	0.112	1.652
Netherlands	564	0.573	0.384	0.104	1.737
Canada	564	0.551	0.378	0.038	1.231
Germany	564	0.270	0.177	0.056	0.651
France	564	0.398	0.336	0.019	1.020
Japan	564	0.637	0.289	0.175	1.285

Figure 1: Buffet indicator values for the seven different countries from 1971 up until 2019.



3.2) Data dependent variables sentiment indices

Data for the sentiment index by Baker and Wurgler will be obtained from the personal page of Jeffrey Wurgler. This data for both the non-orthogonalized as well as the orthogonalized is monthly and available in a period from April 1971 up until December 2018. Furthermore, the data for monthly volatility will be obtained from finance yahoo via the volatility index by the Chicago Board Options Exchange (for the S&P 500 (VIX)), where we take adjusted close price as index value. Summary statistics for these three variables can be seen in table 2, where at first glance the values found for maximum (max.), minimum (min.) and mean seem to vastly differ from those found for the Buffett indicator. Unfortunately, after research and contacting the authors of the paper on international sentiment indices, the data in that paper was unavailable and due to replication limitations, these sentiment indices will be left out of this paper.

Table 2: Summary statistics of both sentiment indices and the volatility index in a period from 1971 up until December 2019 (for both sentiment indices data ends in December 2018 and for the volatility index data starts in February 1986) .

Variable (index)	Observations	Mean	Standard deviation	Minimum	Maximum
Orthogonalized Sentiment	573	-0.273	0.892	-2.422	3.1997
Sentiment	573	-0.035	0.884	-2.316	2.939
Volatility	407	19.959	8.197	7.57	61.41

3.3) Anomaly data

The next variables discussed will be the values for the five anomalies mentioned in the theoretical framework being size, value, momentum, profitability, and investment. For the US data, the personal page of Kenneth French has factor values for all 5 anomalies, all except for the momentum can be taken from their five-factor model, whilst the momentum factor has also been published on its own.

For the other countries outside the US, factor premia will be used, which are obtained from globalfactorpremia.org. This monthly data is only available from July 1990 up until October 2018. A major difference between the data for the US and the other countries in this sample is that for the profitability anomaly for the US there is only profitability based on operating profit, while for the other countries there is also factor values based on the cash-based operating profit, which shows even stronger returns, as could be seen in the theoretical framework and the paper by Hanauer (2020).

Summary statistics for the anomaly returns can be seen in table 3. First of all, one can see that the momentum strategy in Japan yields by far the worst returns, whilst the other countries seem to show similar results for all the values. Canada shows the highest momentum factor whilst the Netherlands shows the most extreme factors for the momentum anomaly.

For the investment anomaly again, Japan really seems to be an outlier with the lowest mean, but this time the difference is a lot smaller compared to the other countries. The means all seem distributed between 0.12 to 0.40, where Canada shows the highest factors. The standard deviations are also equal between 2 and 3.2, even the extreme values for the investment anomaly seem to follow similar values across all the countries.

Furthermore, for the profitability based on operating profit it seems to, again, show similar results for the countries where most countries show some good returns on average except for Japan, where this time the returns are even negative. Extreme values however for all countries seem to show similar values. For this anomaly however Canada slightly shows better results combined with a slightly higher standard deviation than the rest, but the differences are not big.

Next for the size factor we see the returns are not as great as for the other anomalies. Where for two of the seven countries we see a negative mean value. Moreover, it seems as if for the size factor mostly the two North American countries, Canada, and the US, have the best results. In general, the spread between the min. and max. values are quite similar between the countries and when compared to the other anomalies.

The value factor seems to show similar returns across all seven countries with no country really looking out of place. The extreme values and standard deviations for this anomaly also do not really show anything weird.

For the final anomaly. the profitability anomaly based on cash, where data was unavailable for the US, it seems again that Japan has the lowest return mean, this time positive however, and the rest of the countries have similar returns to each other, whilst Canada has the best returns and a slightly higher standard deviation, but nothing out of the ordinary.

Table 3**:** Summary statistics for the monthly momentum, profitability (both cash and not just cash based) and investment anomaly returns in a period from April 1971 up until December 2019 (for countries outside of the US data starts in July 1990).

Nation	**Observations**	**Mean**	**SD**	**Minimum**	**Maximum**
Momentum					
United States	585	1.880	14.72	-107.53	53.38
United Kingdom	340	1.007	4.178	-28.493	11.757
Netherlands	340	0.736	5.771	-34.426	26.278
Canada	340	1.144	4.871	-25.193	19.860
Germany	340	0.950	4.720	-24.151	20.933
France	340	0.600	4.614	-24.435	20.395
Japan	340	0.048	4.463	-23.900	12.843
Profitability					
United States	585	0.261	2.547	-13.320	14.300
United Kingdom	340	0.189	2.208	-7.728	10.021
Netherlands	340	0.125	3.932	-13.180	15.862
Canada	340	0.660	4.216	-11.786	18.872
Germany	340	0.530	2.305	-6.712	10.880
France	340	0.225	2.320	-10.182	9.137
Japan	340	-0.009	1.721	-5.788	7.537
Investment					
United States	585	0.259	2.488	-10.42	14.850
United Kingdom	340	0.319	2.105	-5.326	9.362
Netherlands	340	0.119	3.314	-11.590	12.916
Canada	340	0.126	3.323	-11.930	11.824
Germany	340	0.395	3.054	-9.550	18.149
France	340	0.313	2.529	-7.411	9.33
Japan	340	0.057	2.181	-10.956	7.594
Value					
United States	585	0.300	2.895	-11.120	12.580
United Kingdom	340	0.208	2.995	-13.536	16.516
Netherlands	340	0.427	4.090	-12.594	11.381
Canada	340	0.262	4.254	-27.402	25.640
Germany	340	0.568	3.432	-16.924	20.808
France	340	0.304	3.501	-32.108	11.985
Japan	340	0.437	2.603	-6.2659	7.5364
Size					
United States	585	0.154	2.970	-14.890	18.080
United Kingdom	340	0.089	3.184	-12.576	13.409
Netherlands	340	-0.244	3.314	-9.854	14.476
Canada	340	0.178	2.946	-13.811	13.288
Germany	340	-0.072	2.968	-10.919	9.719
France	340	0.017	2.948	-10.190	10.426
Japan	340	0.083	3.340	-12.342	14.074
Profitability (Cash					
United Kingdom	340	0.275	2.078	-7.638	9.321
Netherlands	340	0.335	3.225	-13.328	10.944
Canada	340	0.630	3.489	-11.407	10.495

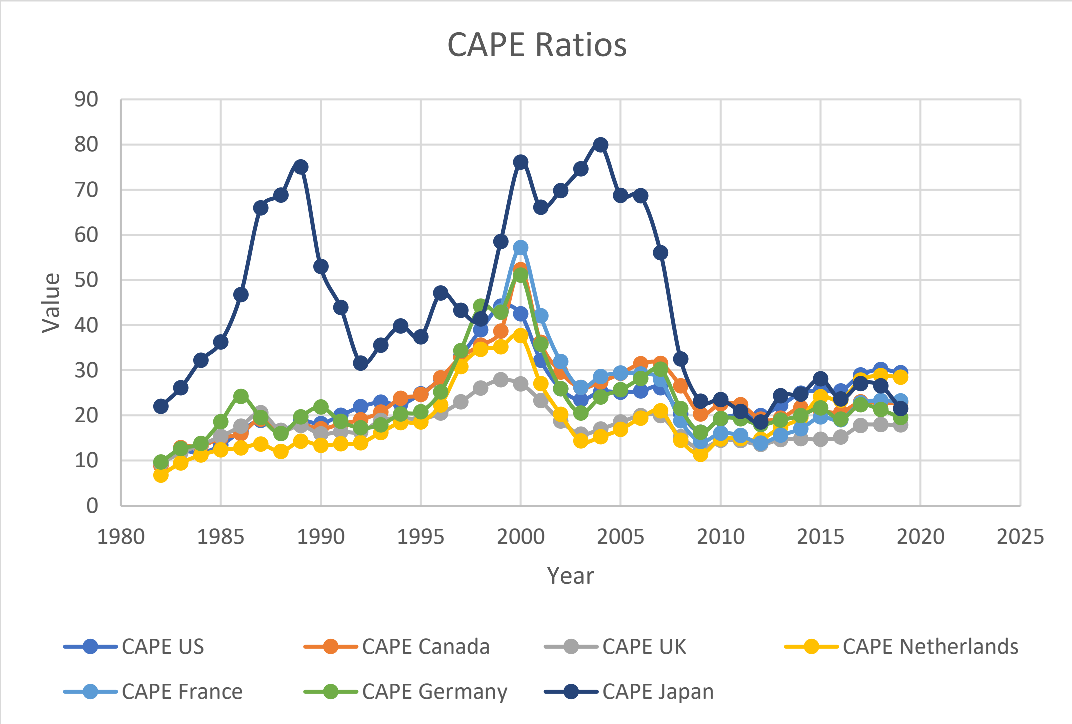
3.4) CAPE data

Lastly monthly data for all CAPE ratios has been obtained from Barclays, who in turn have used data from DataStream and MSCI inc. Summary statistics for the CAPE ratios are available in table 4, in which the Netherlands and the UK have the lowest CAPE ratio, whilst Japan has the highest. On average the CAPE ratio for all countries is similar, except for Japan which really stands out, for both mean and standard deviation. When examining figure 2, the same can be noticed, where pretty much all countries have similar values and follow similar trends, except for Japan, which up until the banking crisis in 2008, has much higher highs and looks a lot more volatile. The last thing to note is the fact that for the French CAPE ratio, there are only 251 observations whilst for the rest there are 456 observations over the testing period. All data for the CAPE is also monthly available for all countries, except France from January 1982 up until December 2019, ,whilst for France data starts in February 1999. It is of importance to note that the US CAPE will not be orthogonalized, as it was in former research, due to comparability with the other countries, since the orthogonalization process cannot be performed for these other countries, due to unavailability of this data for the other countries. The orthogonalization factors used for the US are unfortunately not available for the other countries in this sample, even after reaching out to multiple data suppliers and authors.

Table 4**:** Summary statistics of CAPE ratios for 7 countries in a period stretching from January 1982 up until December 2019.

Nation	**Observations**	**Mean**	**Standard deviation**	**Minimum**	**Maximum**
United States	456	23.692	7.792	8.060	47.130
United Kingdom	456	17.631	4.095	8.770	28.500
Netherlands	456	18.978	7.666	6.300	40.010
Canada	456	23.803	8.380	8.360	60.620
Germany	456	23.056	8.794	9.050	58.950
France	251	25.396	10.974	11.410	61.260
Japan	456	43.652	19.891	16.600	90.920

Figure 2: Graph showing the CAPE ratios for 7 countries in a period stretching from January 1982 up until December 2019.



3.5) Dickey-Fuller test

For further testing, due to the nature of all the indexes, the CAPE ratio, and the Buffett indicator ,there might be underlying problems with stationarity and unit root, which could cause problems for interpretation by creating spurious regressions and correlations, which could cause false claims being made. To use these variables for testing first an augmented Dickey-Fuller test with a drift term and lag will be run, with a following formula:

$$\Delta y_i = a_0 + \beta y_{i-1} + u_i. \qquad (1)$$

Where we reject the null if β is not equal to 0, which would mean a unit root problem is present. If a unit root problem is present, a new variable regarding that variable will be created, taking the first difference of that variable, and using this for further testing.

3.6) Correlation and regressions (Sentiment proxy tests)

Next, to check whether the US Buffett indicator and CAPE ratio follow a similar trend to the dependent variables, both sentiment indices and the volatility index, a correlation test will be run. The correlation test run in this paper will be the Pearson correlation coefficient test, which is a test for linear correlation between two data sets. It is the covariance of the two variables divided by the product of their standard deviations, where the closer the number is to 1 or -1 the more it is correlated. The formula used looks like this:

$$\rho(x, y) = \Sigma[(xi - \bar{x}) * (yi - \bar{y})] / (\sigma x * \sigma y) \quad (2)$$

Where $\bar{x}$ is the mean of the US Buffett indicator or CAPE ratio and xi is the observed value of the US Buffett indicator or CAPE ratio, whilst σx is the standard deviation of the US Buffett indicator or CAPE ratio. Where y denotes the same values but for the dependent variables, being both the sentiment indices and the volatility index.

After which first the US Buffett indicator and US Shiller CAPE ratio will be regressed on the same dependent variables. For this step, a standard OLS (Ordinary Least Squares) regression will be used, following this form:

$$Y_{t} = \beta 0 + \beta 1 X1_{t-1} \quad (3)$$

Where Y_t is the value of the dependent variables, being both the sentiment indices and the volatility index. β_0 is the constant, β_1 is the coefficient showing the linear relation between the Buffett indicator or Cape ratio with the dependent variables and where $X_{1\ t-1}$ is the independent variable, being either the Buffett indicator or Cape ratio. Where, if any link can be found by any of these methods, both hypothesis 1a and 1b cannot be rejected and the Buffett indicator can be seen as a sentiment proxy. However, if the Buffett indicator is not found to be a sentiment proxy, it can still show predictive power for hypothesis 2a

3.7) T-test (Investment strategy)

For hypothesis 2a a dummy variable will be created to assign high or low Buffett indicator values to certain months for all seven countries in the sample. A high Buffett indicator value month is defined as a month where the previous value of the Buffett indicator in question is above the median value of the entire sample, a rolling median did not show massive changes in data, which is why in this paper, I have chosen to use the median over the entire sample for simplicity in replication purposes. For this test to determine high or low values, the normal Buffett indicator values are used, since the dummy variable being created is not subject to the unit root problem. Furthermore, due to future information being used to calculate this median, results concerning future returns should be taken with a pinch of salt. Then a two-sample t-test using groups will determine whether a high Buffett indicator has influence on anomaly factor values and whether the Buffett indicator can be used for an investment strategy. This T-test will look as follows:

$$T = \frac{X1 - X2}{(Sp * \sqrt{\frac{1}{n1} + \frac{1}{n2}}} \tag{4}$$

Where X_1 is the mean of the high-valued months and X_2 is the mean of the low valued months. SP is the pooled standard deviation and $n_1$and n_2 are the observations per subcategory of the Buffett indicator or Shiller CAPE ratio.

3.8) Predictive power regression

To further test the predictive power of the Buffett indicator on the anomaly factor values regressions will be run where the dependent variables will be the anomaly factor values:

$$Y_t = \beta 0 + \beta 1 X_{t-1} \quad (5)$$

Where Y_t is the value of the dependent variables, being both the factor values for the different anomalies and the different countries. β_0 is the constant, β_1 is the coefficient showing the linear relation between the Buffett indicator, adjusted for the unit root problem, with the dependent variables and where X_{t-1} is the independent variable, being the adjusted Buffett indicator. For hypothesis 2b a similar method will be conducted for the CAPE ratio, where the only difference is that this time X_{t-1} is the value of the CAPE.

3.9) Equal weighted

Furthermore, to test for a general ability to predict factor values, an equal weighted average regression will be run for both the Buffett indicator and CAPE ratio for each specific anomaly. Where again a similar method as seen in formula 5 will be constructed, where this time Y_t is the equal weighted value of the anomaly factors across the different countries in the sample and X_{t-1} is the equal weighted value of either the Buffett indicator or CAPE ratio across the different countries.

Just like the regressions being run, similar T-tests to the ones run for the separate countries will be run for the equal weighted returns and Buffett indicator and Shiller CAPE.

Lastly for the equal weighted returns and Buffett indicator and Shiller CAPE ratio, Sharpe ratios and turnover ratios will be calculated. Sharpe ratios are a common test to judge the profitability of a certain investment strategy or portfolio, with most investors accepting strategies or portfolios above the value of 1 as profitable. Sharpe ratios will be calculated as follows :

$$Sp = \frac{Rt - Rf}{SD} \quad (6$$

Where Sp is the Sharpe ratio, Rt is the return of the strategy, Rf is the risk-free rate and SD is the standard deviation of the excess returns following the return of the strategy minus the risk-free rate. In this case the risk-free rate has been taken from the personal website of Kenneth French.

For the market the Sharpe ratio, calculated using the data from the Kenneth French data library, for the period from 1970 up until December 2020, is 0.322, which will be used as a benchmark throughout this paper.

The turnover ratio is the number of times during the entire period of testing the Buffett indicator or Shiller CAPE ratio switches form sign, meaning how many times does it switch from high median to low median. This will give an idea about the transaction costs endured during the trading strategy.

4) Results:

4.1) Results of trend testing

First, Table 5 shows the results of the augmented Dickey-Fuller for all variables that might experience non-stationarity and unit root problems. First it can be seen that none of the dependent variables for hypothesis 1 experience any unit root problems at a significance level of 0.05, so normal values for these variables can be used for testing. For the Buffett indicators it seems all the countries in the sample seem to experience unit root problems. Only Japan would not show signs of a unit root problem at a significance level of 0.10, but this is not a common level to use and therefore shall be disregarded. For further testing of the Buffett indicator, a new variable must be created measuring the first difference between the values of the Buffett indicator and its one-month lagged values, as mentioned. For the CAPE ratio it is unfortunately not that easy, where all countries show similar results. In this case the Netherlands and France both show signs of a unit root problem. Due to the similarity of the CAPE ratios, it seems not logical to use different approaches for different countries and for complete safety a similar process around the first difference will also be applied to the CAPE ratios, to guarantee the most realistic results. However, due to uncertainty the same tests will also be run for the non-first difference values of the CAPE ratio, to not exclude these values for the countries not experiencing unit root problems with their CAPE ratio. Lastly it can be seen that, for the equal weighted Buffett indicator, there seems to be a problem with unit root as well, therefore a first difference variable will be used in its place for the regressions run on the equal-weighted average Buffett indicator.

Similar results are not seen for the equal-weighted average of the CAPE ratio, where a unit room problem does not seem to cause an issue, but for comparability reasons the first difference of the equal weighted average of the PE will also be included in this paper.

Table 5: Results of the augmented Dicky fuller test, with drift and one lag, for the Buffett indicators, sentiment indices, volatility index, CAPE ratios and equal weighted averages.

	Z-score		Z-score		Z-score
Dependent Variables		**Buffett Indicator (Countries)**		**CAPE Ratio (Countries)**	
Orthogonalized Sentiment index	-2.403 ***	United States	-0.369	United States	-1.874 **
Sentiment index	-2.219**	United Kingdom	-1.233	United Kingdom	-2.632 ***
Volatility index	-5.604***	Netherlands	-1.196	Netherlands	-1.544 *
		Canada	-0.962	Canada	-2.235 **
		Germany	-1.078	Germany	-2.288 **
		France	-0.600	France	-1.590 *
		Japan	-1.425 *	Japan	-1.723 **
		Equal Weighted Average		**Equal Weighted Average**	
		Buffett indicator	-0.818	CAPE ratio	-2.110 **

*** significant at the 0.01 level
** significant at the 0.05 level
* significant at the 0.1 level

Table 6 shows all results concerning the testing run for hypothesis 1. Where first we can see that the US Buffett indicator does not show any significant correlation with the two sentiment indices. It does however show strong significant negative correlation with the volatility index, which was expected, since high volatility values usually occur when market value is low. The regression results show the same, where again the US Buffett indicator only seems to significantly affect the volatility. As mentioned before, the first difference variable of the US Buffett indicator has been used to conduct both these tests, due to stationarity problems with the normal values of the US Buffett indicator.

Furthermore, for the US CAPE ratio we see for the level variables that the results in general are a lot less significant than for the first difference, which confirms the idea that the first differences of all CAPE ratios will provide better results for interpretation. When looking at the first difference US CAPE ratio, one can see that it significantly correlates to both the sentiment indices as well as the volatility index. The same results follow from the Beta coefficients from the run regressions. Just like for the Buffett indicator the signs are negative, meaning there is a negative relation between the US

CAPE ratio and the dependent variables, which for the sentiment index is the opposite of what was expected, but for the volatility follows the expectations behind hypothesis 1.

A visual representation for the US Buffett indicator and the US CAPE ratio and the dependent variables, in figure 3 and 4, follows the same conclusions, where no real similar trend can be seen between these variables and the dependent variables.

In conclusion one can reject hypothesis 1a since there is no significant relation between the US Buffett indicator and the sentiment proxies. There is a significant relation between the CAPE ratio and the sentiment proxies. However, expectations were that the sentiment indexes and the Shiller CAPE as well as the Buffett indicator follow similar positive trends, which in this case is not present since the significant relation between the CAPE ratio and the sentiment indices is negative. Hypothesis 1b cannot be rejected since both the US CAPE ratio and the US Buffett indicator show a significant negative relation with the volatility index and therefore show a similar trend to the volatility index as a sentiment proxy.

Table 6**:** Table showing the results of the correlation and regression run on the US Buffett indicator and the independent variables, orthogonalized and non-orthogonalized sentiment index (1971-2018) and the volatility index (1986-2019)

	Orthogonalized Sentiment Index	**Sentiment Index**	**Volatility Index (VXO)**
Buffett Indicator (First Difference)			
Correlation	-0.0078	-0.0314	-0.3106***
Beta coefficient	-0.237	-0.947	-78.114***
CAPE ratio			
Correlation	0.1121**	-0.0145	0.0296
Beta coefficient	0.009**	-0.001	0.034
CAPE ratio (First Difference)			
Correlation	-0.1039**	-0.0958**	-0.2376***
Beta coefficient	-0.062**	-0.064**	-1.762***

*** significant at the 0.01 level
** significant at the 0.05 level
* significant at the 0.1 level

Figure 3: US Buffett indicator and the dependent variables.

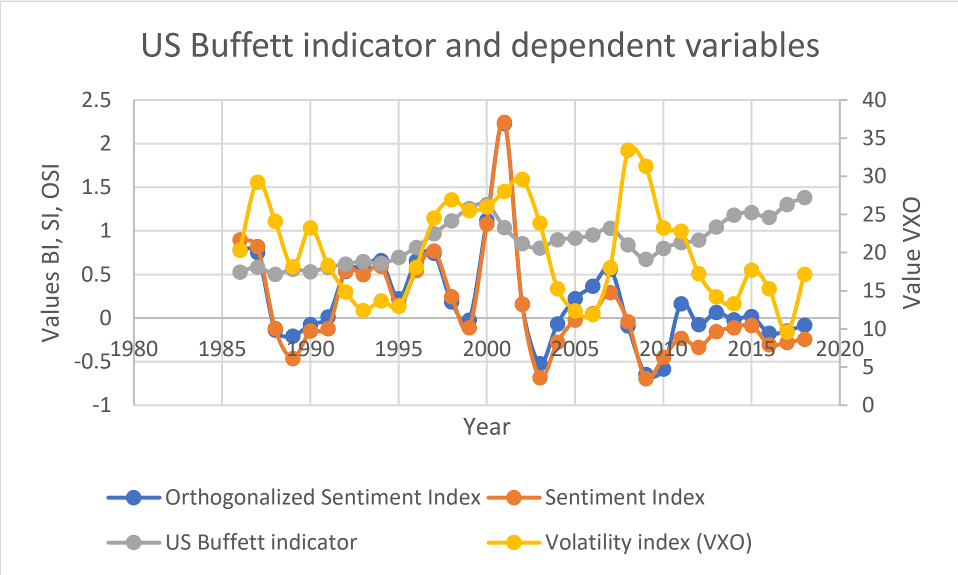
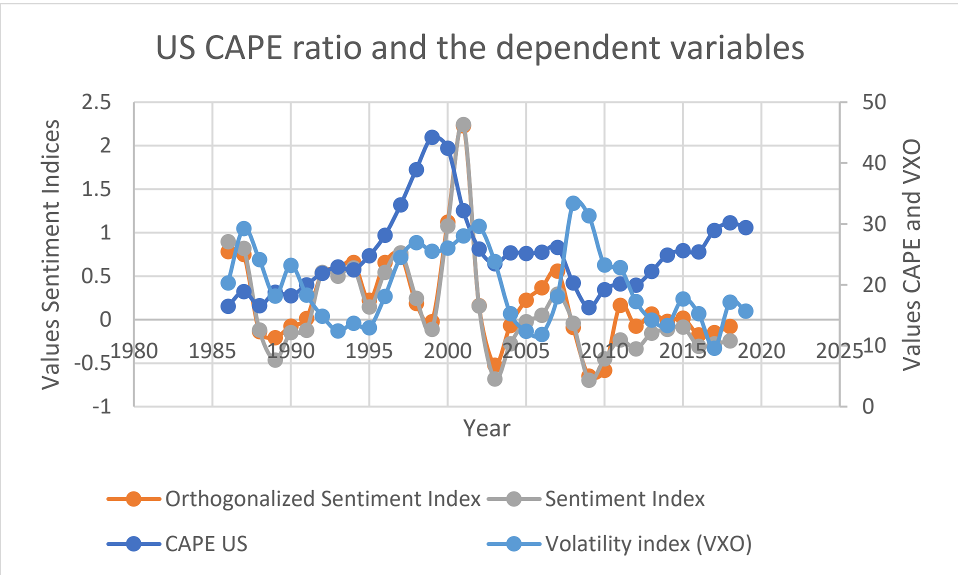


Figure 4: US Buffett indicator and the dependent variables.



4.2) Results Buffett indicator

For hypothesis 2a two tests have been run. Table 7 shows the results of the two-sample T-test using groups, being differentiated between high and low Buffett indicator values. We can see that for most countries and most anomalies the difference between high and low months is positive, however in general, there is a lack of significant results being shown in table 7. The only significant positive difference that can be seen in table 7, between high and low value months, is for the size anomaly in the Netherlands. This conclusion for the Netherlands follows nicely with the findings by Miller (1977) and Stambaugh, Yu and Yuan (2011), where periods of overpricing showed better results for anomalies, since periods of high Buffett indicator values are a sign of overpricing (Buffett,1999).

When further looking at the results of table 7, which shows the results of the regressions of the Buffett indicators on the anomalies per country, we see more significant results. First of all, for the US we see a significant positive effect of the Buffett indicator on the size anomaly at a significance level of 0.05, again showing the same conclusions as drawn by Stambaugh, Yu and Yuan (2011) of the positive relation between indicators of periods of overpricing and anomaly returns . For the same anomaly, being size, in the UK and France we see a significant effect at a level of 0.01. For the rest of the countries, there is no similar significance shown in table 7 on the size anomaly. Two more significant results can be seen for the investment anomaly and the UK and Japan Buffett indicators, where both show a significant negative relation between the factor values of the investment anomaly and the mentioned Buffett indicators, which is the first significant negative relation shown between the Buffett indicator and anomaly returns. The negative relation between the investment anomaly and Buffett indicator does not seem to follow the conclusions of Stambaugh et al.(2011), since in this case periods of overpricing do not seem to show better anomaly returns, which is first evidence that the Buffett indicator may not show as strong as an effect on anomalies as for example the Baker and Wurgler sentiment index did in the paper by Stambaugh et al. (2011) Lastly for France table 7 shows a significant negative relation between the French Buffett indicator and the factor values of the French cash-based profitability anomaly, which again shows the opposite as the expected relation following the paper by Stambaugh et al (2011).

Table 8 shows the results of the equal weighted averages of the Buffett indicator on the equal weighted averages of the anomalies. In table 8 we first see that only the profitability anomaly has a positive significant difference for high and low values of the Buffett indicator and that in general the differences are not even always positive, which shows that, compared to the Baker and Wurgler sentiment proxy, the Buffett indicator in general, looking at the 7 countries in this sample, shows weaker predictive power of overpricing. Table 8 further shows a completely different result, where only for the size anomaly, there is a positive significant relation between the Buffett indicator and an anomaly, which does follow the findings in Stambaugh et al. (2011), showing a positive relation between overpricing proxies and anomalies.

In general, the relation between the Buffett indicator and the anomalies is not often there and a difference between high and low valued Buffett indicator months, most of the time, does not show the expected positive significant difference following the paper by Stambaugh et al.(2011) in factor values for the anomalies. The regressions seem to follow suit where only for certain countries and certain anomalies there seems to be a significant effect, whilst also not always being positive.

When further looking at table 9, which shows the Sharpe ratios of the trading strategies using the equal weighted returns on the anomalies and the turnover, one can see that the trading strategies do not show very good results compared to the market benchmark of 0.322, with exception for the cash-based profitability anomaly, which shows a value of 0.624, massively outperforming the market. Further looking at the turnover ratio being so low one can disregard any transaction costs that might be endured during the execution of said trading strategy.

Table 7**:** Table showing the T-test scores and Beta coefficients for the Buffett indicators on the monthly anomaly return factors (1971-2019 for the US and UK, 1973-2019 other countries)

Country	High	Low	Difference HIgh-LOW (t-value)	Beta coefficient
USA				
Momentum	0.837	0.466	0.371 (1.0374)	10.800* (6.021)
Profitability	0.432	0.143	0.289* (1.5516)	-4.763 (3.134)
Investment	0.185	0.394	-0.208 (-1.294)	1.293 (2.719)
Size	0.053	0.254	-0.201 (-0.818	9.926** (4.123)
Value	0.201	0.398	-0.197 (-0.823)	3.629 (4.037)
UK				
Momentum	1.048	0.831	0.217 (0.371)	-1.928 (4.037)
Profitability	0.221	0.051	0.169 (0.549)	-3.462 (2.126)
Size	0.145	-0.156	0.301 (0.677)	11.960*** (3.008)
Value	0.189	0.294	-0.105 (-0.251)	0.459 (2.895)
Investment	0.322	0.307	0.015 (0.049)	-4.176** (2.021)
Cash-based profitability	0.255	0.364	-0.109 (-0.375)	-2.936 (2.002)
Netherlands				
Momentum	0.721	0.790	-0.069 (-0.091)	7.301 (5.893)
Profitability	0.234	-0.276	0.510 (0.982)	-5.370 (4.014)
Size	-0.054	-0.935	0.881 ** (2.021)	6.593* (3.373)
Value	0.316	0.834	-0.518 (-0.959)	0.427 (4.185)
Investment	0.225	-0.269	0.494 (1.129)	0.068 (3.392)
Cash-based profitability	0.401	0.069	0.338 (0.794)	-1.394 (3.300)
Canada				
Momentum	1.064	1.437	-0.373 (-0.579)	0.965 (5.178)
Profitability	0.843	-0.012	0.855 * (1.539)	2.565 (4.479)
Size	0.121	0.388	-0.267 (-0.687)	3.870 (3.124)
Value	0.416	-0.300	0.717 (1.277)	0.839 (4.522)
Investment	0.247	-0.318	0.565 * (1.324)	1.858 (3.435)
Cash-based profitability	0.635	0.609	0.026 (0.057)	-2.432 (3.707)
Germany				
Momentum	0.982	0.835	0.147 (0.236)	8.009 (10.56)
Profitability	0.549	0.458	0.091 (0.301)	1.482 (5.159)
Size	0.004	-0.347	0.351 (0.900)	9.740 (6.623)
Value	0.588	0.497	0.091 (0.202)	-7.135 (7.675)
Investment	0.474	0.112	0.363 (0.903)	-11.910* (6.807)
Cash-based profitability	0.491	0.346	0.145 (0.491)	3.496 (5.016)
France				
Momentum	0.544	0.806	-0.262 (-0.430)	6.685 (6.381)
Profitability	0.141	0.530	-0.389 (-1.272)	-1.940 (3.211)
Size	0.081	-0.215	0.296 (0.760)	12.240*** (4.029)
Value	0.432	-0.163	0.595 * (1.289)	1.111 (4.849)
Investment	0.386	0.045	0.341 (1.020)	-5.215 (3.491)
Cash-based profitability	0.365	0.600	-0.235 (-0.762)	-8.618*** (3.202)
Japan				
Momentum	0.272	-0.363	0.6335 (1.254)	-2.290 (5.321)
Profitability	-0.016	0.002	-0.0175 (-0.090)	3.377* (2.045)
Size	-0.007	0.248	-0.255 (-0.671)	3.024 (3.980)
Value	0.539	0.249	0.290 (0.981)	-0.754 (3.104)
Investment	0.089	-0.001	0.091 (0.367)	-5.984** (2.581)
Cash-based profitability	0.098	0.124	-0.026 (0.108)	3.069 (2.518)

*** significant at the 0.01 level
** significant at the 0.05 level
* significant at the 0.1 level

Table 8: T-test and regression (Beta coefficient) results of the equal weighted averages of the Buffett indicator on the equal weighted average of the anomalies (1971-2019)

Anomaly	High	Low	Buffett Indicator (first difference) (standard error)	Difference HIgh-LOW (t-value)
Momentum	0.870	0.659	3.222 (4.700)	0.211 (0.688)
Profitability	0.375	0.147	-1.031 (1.859)	0.228 ** (1.880)
Investment	0.229	0.350	9.640*** (2.958)	-0.121 (-0.899)
Size	-0.053	0.212	-2.022 (3.210)	-0.264 (-1.359)
Value	0.324	0.426	-3.355 (2.066)	-0.102 (-0.489)
Cash-based profitability	0.391	0.303	-2.102 (1.825)	0.089 (0.582)

*** significant at the 0.01 level
** significant at the 0.05 level
* significant at the 0.1 level

Table 9 : Sharpe and turnover ratios of the equal weighted trading strategy using the Buffett indicator (1971-2019)

Anomaly	Sharpe ratio	Turnover ratio
Momentum	0.091	2.22%
Profitability	0.255	2.22%
Size	-0.201	2.22%
Value	-0.076	2.22%
Investment	-0.137	2.22%
Cash-based profitability	0.629*	2.22%

* Outperforms the market (0.322)

4.3) Results CAPE ratio

The results of the tests for the CAPE ratio can be found in table 10, 11 and 12. Table 10 shows the results of the two sample T-test using groups, where the groups in this case are months with high and low valued CAPE ratios. We see different results compared to the Buffett indicator with more significant observations for different countries and anomalies. First of all, there appears to be a large significant positive difference between high and low CAPE months on the momentum anomaly for the US, UK, Germany, and Japan, which follows the findings of Stambaugh et al. (2011) and shows that the CAPE is a better predictor of overpricing for these countries when it comes to the momentum anomaly. For the Netherlands and Canada there seems to also be a large positive difference, but without any statistical significance. For France however there seems to be a negative difference between high and low months, without any statistical significance. In general, all the results from France seem to be non-significant, due to data limitations and the lower number of observations.

Further there seems to only be any significance for the value anomaly for both the UK and Japan CAPE ratio, this difference is positive which shows that periods following high CAPE values, which are a sign of overpricing, show better results for the value anomaly which follows the findings of Stambaugh et al.(2011) . For all other anomalies and countries there seem to be no significant results at the 0.05 level. There are some observations that show a significant positive difference at the 0.1 level, but this is not a standard upheld criteria to show significance in economic research. Just like for the Buffett indicator the difference between high and low months is not always positive and for the CAPE ratio there are even some non-positive significant results for example for the investment anomaly and the Japanese CAPE ratio, which completely contradict expectations, since overpricing is usually paired with stronger anomaly returns following Stambaugh et al. (2011).

Table 10 further shows that the CAPE ratio shows more predictive power than the Buffett indicator. In this paper as mentioned earlier we will focus mainly on the first difference of the CAPE due to stationarity problems, therefore focus will lay on the rightest column in table 10. The size anomaly shows the strongest relation with the CAPE ratio, since it shows a significant positive relation for the US, the UK, the Netherlands, Canada, Germany, and France, only for Japan it does not show a significant relation, which is in line with the findings of Stambaugh et al. (2011), where again overpricing proxies were found to also be good predictors of anomaly returns in general. Furthermore, for the UK and Germany table 10 shows a significant relation for both the profitability and cash-based profitability anomaly, but for the UK this is a negative relation, whilst for Germany it is a positive beta coefficient, which is quite peculiar since the German and UK Shiller CAPE ratio seem to follow similar patterns. When comparing these two outcomes, the positive relation is in line with the findings by Stambaugh et al. (2011), whilst the findings for the UK are of the complete opposite. Lastly, the Japanese CAPE ratio does show two significant relations, a positive relation with the cash-based profitability anomaly, which is in line with the expectations and a negative relation with the investment anomaly, which is completely the opposite with expectations. Apart from all these significant results, most of the coefficients are negative, showing non-significant negative relation between most CAPE values and anomalies, which was not expected when looking at the properties of the CAPE ratio and its predictive power of overpricing, linked with the paper of Stambaugh et al.(2011), however these results are non-significant, which makes interpreting them not that is interesting.

When further examining table 11, one can see that, just like for the equal weighted average of the Buffett indicator, overall results for the equal weighted average of the CAPE ratios seem not significant. When looking more in depth into table 11, we see only a significant positive difference between high and low months for the momentum anomaly, which is in line of expectation following research by Stambaugh et al.(2011). For the other anomalies there is no significant difference, nor are the differences even positive all the time, which is quite unexpected, since for the individual countries were a few more significant relations and in general, following Stambaugh et al. (2011), results were expected to be positive. Table 11 furthermore does show the best results so far, with significant positive coefficients for the momentum, investment, and value anomaly, which again means results are in line with the findings of Stambaugh et al. (2011). Showing that for half the investigated anomalies, the CAPE ratio shows strong predictive power. However, this alone is not enough evidence to not reject hypothesis 2b since this makes claims on predictive power on all anomalies. Again, for continuity purposes, this paper will solely focus on the first difference of the CAPE, even though the equal weighted average does not show strong unit root problems. When looking at the non-first difference CAPE ratio variables, table 11 shows fewer positive results, where only momentum shows a significant positive relation.

When then further looking into table 12, which shows the Sharpe ratios and turnover percentage of the anomaly-based trading strategy involving Shiller's CAPE ratio as predictor for overpricing, one can see that just like for the Buffett indicator, results are bad and even worse than for the Buffett indicator, with a highest Sharpe ratio of 0.281, which means none of the strategies outperform the market . When further examining the turnover percentage we can see this percentage, just like for the Buffett indicator, is low, but slightly higher, however it is low enough to not cause any issues if one were to implement the strategies.

Table 10 : Table showing the T-test scores and Beta coefficients for the CAPE ratios on the anomaly factors (1999-2019 for France and 1982-2019 for the other countries)

Country	High	Low	Difference HIgh-LOW (t-value)	Beta coefficient	First difference Beta Coefficient
USA					
Momentum	0.908	0.175	0.733 ** (1.779)	0.054** (0.027)	-0.107 (0.197)
Profitability	0.449	0.263	0.185 (0.830)	0.0007 (0.014)	-0.166 (0.106)
Investment	0.149	0.404	-0.255 (-1.378)	0.006 (0.012)	0.0165 (0.0876)
Size	0.078	0.052	0.026 (0.098)	-0.002 (0.017)	0.358*** (0.127)
Value	0.226	0.260	-0.033 (-0.123)	0.0003 (0.017)	0.110 (0.128)
UK					
Momentum	1.360	0.607	0.752 ** (1.662)	0.045 (0.055)	-0.209 (0.309)
Profitability	0.0002	0.404	-0.404 (-1.689)	-0.035 (0.029)	-0.417** (0.162)
Size	0.040	0.144	-0.104 (-0.299)	0.003(0.042)	1.437*** (0.222)
Value	0.478	-0.100	0.578 ** (1.782)	0.065 (0.039)	0.112 (0.221)
Investment	0.238	0.412	-0.174 (-0.758)	0.010 (0.028)	-0.057(0.155)
Cash-based profitability	0.170	0.394	-0.224 (-0.9924)	-0.011 (0.028)	-0.374** (0.152)
Netherlands					
Momentum	1.031	0.248	0.783 (1.212)	0.055 (0.042)	-0.130 (0.257)
Profitability	0.232	-0.053	0.286 (0.649)	0.040 (0.029)	0.080 (0.175)
Size	-0.157	-0.387	0.229 (0.618)	0.012 (0.024)	0.470*** (0.146)
Value	0.411	0.454	-0.043 (-0.093)	-0.012 (0.030)	-0.155 (0.182)
Investment	0.283	-0.154	0.437 (1.178)	0.016 (0.024)	-0.078 (0.148)
Cash-based profitability	0.305	0.384	-0.079 (-0.218)	0.021 (0.024)	-0.014 (0.144)
Canada					
Momentum	1.432	0.641	0.791 * (1.444)	0.057* (0.034)	0.128 (0.182)
Profitability	0.896	0.247	0.649 * (1.368)	0.057* (0.029)	-0.181 (0.157)
Size	0.091	0.329	-0.328 (-0.716)	-0.020(0.020)	0.368*** (0.108)
Value	0.412	0.0002	0.412 (0.860)	0.060** (0.029)	-0.099 (0.159)
Investment	0.326	-0.222	0.549 * (1.510)	0.031 (0.022)	-0.121 (0.121)
Cash-based profitability	0.778	0.372	0.406 (1.034)	-0.005 (0.024)	-0.205 (0.130)
Germany					
Momentum	1.560	0.088	1.472 *** (2.863)	0.055* (0.028)	-0.001 (0.146)
Profitability	0.577	0.462	0.115 (0.455)	0.011 (0.014)	0.144** (0.071)
Size	0.007	-0.184	0.192 (0.586)	-0.029 (0.018)	0.190** (0.092)
Value	0.605	0.516	0.089 (0.236)	0.039* (0.021)	-0.151 (0.106)
Investment	0.597	0.111	0.485 * (1.446)	0.063*** (0.018)	-0.043 (0.095)
Cash-based profitability	0.457	0.462	-0.005 (-0.022)	0.023* (0.013)	0.143** (0.069)
France					
Momentum	0.544	0.806	-0.262 (-0.430)	0.032 (0.030)	-0.044 (0.228)
Profitability	0.152	0.295	-0.142 (-0.449)	0.082 (0.107)	0.082 (0.107)
Size	0.277	0.209	0.068 (0.1947)	0.010 (0.016)	0.302** (0.117)
Value	0.641	0.128	0.513 (1.002)	0.030 (0.023)	-0.216 (0.173)
Investment	0.588	0.230	0.358 (1.043)	0.042*** (0.015)	-0.117 (0.117)
Cash-based profitability	0.304	0.584	-0.280 (-0.828)	0.008 (0.015)	-0.065 (0.115)
Japan					
Momentum	0.505	-0.448	0.953 ** (1.976)	0.017 (0.012)	-0.050 (0.069)
Profitability	-0.027	0.010	-0.373 (-0.199)	-0.005 (0.005)	0.033 (0.027)
Size	-0.101	0.283	-0.384 (-1.061)	-0.003 (0.009)	0.058 (0.052)
Value	0.708	0.142	0.567 ** (2.014)	0.023*** (0.007)	-0.010 (0.041)
Investment	-0.147	0.279	-0.425 (-1.803)	-0.004(0.006)	-0.071** (0.034)
Cash-based profitability	0.114	0.099	0.016 (0.068)	-0.007 (0.006)	0.083** (0.033)

*** significant at the 0.01 level
** significant at the 0.05 level
* significant at the 0.1 level

Table 11: T-test results and Beta coefficients of the equal weighted averages of the CAPE ratio on the equal weighted average of the anomalies.

ce

Anomaly	High	Low	CAPE ratio (level) (standard error)	CAPE ratio (First-difference) (standard error)	Difference HIgh-LOW (t-value)
Momentum	0.988	0.382	0.043** (0.021)	0.059** (0.025)	0.606 ** (1.794)
Profitability	0.268	0.377	0.008 (0.008)	0.0129 (0.009)	-0.109 (-0.836)
Investment	0.327	0.220	-0.010 (0.012)	-0.002 (0.013)	0.107 (0.729)
Size	-0.060	0.003	0.0253* (0.014)	0.048*** (0.015)	-0.063 (-0.338
Value	0.419	0.246	0.013 (0.009)	0.025** (0.010)	0.173 (0.766)
Cash-based profitability	0.338	0.408	0.009 (0.008)	0.008 (0.008)	-0.070 (-0.536)

*** significant at the 0.01 level
** significant at the 0.05 level
* significant at the 0.1 level

Table 12 : Sharpe and turnover ratios of the equal weighted trading strategy using the Shiller CAPE Ratio (1971-2019)

Anomaly	Sharpe ratio	Turnover ratio
Momentum	0.282	4.44%
Profitability	-0.142	4.44%
Size	-0.060	4.44%
Value	0.118	4.44%
Investment	0.109	4.44%
Cash-based profitability	-0.012	4.44%

* Outperforms the market (0.322)

So just like for the Buffett indicator, although results for the CAPE ratio show more predictive power, the predictive power of the CAPE is still not always positive, nor is it always significant. Again, there are a few anomalies for a few countries that do show positive significant relation between high and low months and for regressions, but the evidence leads me to reject hypothesis 2b.

In conclusion both hypothesis 2a and 2b can therefore be rejected.

5) Conclusion:

5.1) Conclusion

This paper set out to investigate the predictive power of the Buffett indicator on various anomalies, momentum, profitability, size, value, investing and cash-based profitability, and how it would hold up as a sentiment proxy, by comparing it to established proxies such as the sentiment index by Baker and Wurgler (2006). This paper further extended this research by including a similar test for the CAPE ratio of Shiller (1996) to test where such indicators in general stand in predicting and being a part of sentiment proxies. Data has been gathered over various periods, starting in 1971 up until 2019, and for various countries, the US, the UK, the Netherlands, Germany, France, Canada, and Japan. Anomalies used in this paper are well known and are key components of one of the most elaborate asset pricing models in the current financial world, namely the six-factor model by Fama and French (2018). This paper made use of various tests such as a two-sample T-test and an OLS regression.

This paper showed no real link between the established Baker and Wurgler sentiment index and the US Buffett indicator. However, there was a significant negative relation between the VIX (volatility index by the Chicago Board of Options Exchange) and the US Buffett indicator, which was expected, since in general low volatility is paired with overpricing and high values of the US Buffett indicator follow the opposite trend, where high values are paired with overpricing. For the CAPE there was a negative relation between both the sentiment indices and the VIX, which was not fully expected since high sentiment is usually paired with overpricing as are high values of the Shiller CAPE. Therefore, one can conclude that there is no similar trend between the Baker and Wurgler sentiment index and the US Shiller CAPE, but there is a similar trend as a sentiment proxy for the US CAPE compared to the VIX. One can therefore conclude that both the Buffett indicator and the Shiller CAPE ratio do have properties of certain sentiment proxies such as volatility, but do not have the same properties as more complicated sentiment proxies such as the Baker and Wurgler sentiment index from 2006 and 2007.

Further the predictive power of both the Buffett indicator and CAPE ratio on anomalies in different countries was examined. No strong predictive power for all anomalies was found for neither the Buffett indicator nor the CAPE ratio. First, the Buffett indicator only provided some strongly significant positive relation for the size anomaly for the UK and France, which is in line with the expectations following the papers by Miller (1977) and Stambaugh et al (2011) and the properties of the Buffett indicator and its link to overpricing. Further the Buffett indicator showed a negative relation for the investment anomaly for Japan and the cash-based profitability anomaly for France, which in turn was therefore not expected. When taking a further look into the equal weighted averages, only a strong relation with the size anomaly sticks out for the Buffett indicator. Which means there is not enough proof to conclude that there is a strong relation with anomalies in general for the Buffett indicator.

When next looking at the Shiller CAPE ratio a similar conclusion can be drawn, where again not enough proof was found to conclude there being strong predictive power on anomalies for the CAPE ratio. For all countries except Japan, the size anomaly shows a strong positive significant relation with the CAPE ratio for the specific country, which again was expected following the paper by Miller (1977) and Stambaugh et al. (2011) and the Shiller CAPE's ability to judge overpricing in the market. Some other anomalies in some of the countries also show some positive significant relation between the CAPE ratio and the anomaly, for example the cash-based profitability anomaly for Germany. However just like for the Buffett indicator, the difference and the relation are not always positive. For example, looking at the Beta for the investment anomaly of Japan one can see a negative significant relation between the Japanese factor of this anomaly and the Japanese CAPE ratio, which contradicts the findings of the paper by Stambaugh et al. (2011) concerning the link between overpricing proxies and anomaly returns. However, it might also mean that the Shiller CAPE ratio is just not as good of an overpricing proxy as the sentiment index used in the paper by Stambaugh et al. (2011) is. So just like for the Buffett indicator one must conclude that there is not enough evidence assume strong predictive power for the CAPE ratio, even though it does show slightly better results, especially when looking at the coefficients of the first difference of the equal weighted average of the CAPE ratios.

Furthermore, looking at the results of the T-test and the following Sharpe ratios, one cannot conclude that a trading strategy, where investing in the anomalies when the previous month's value of either the Buffett indicator or CAPE were above the median, yields good results. First, not even for all anomalies or in all countries did it yield positive results in general and even when it did these results were mostly non-significant. For the Buffett indicator only, the Netherlands had a significant positive result for the size anomaly. For the CAPE ratio there were slightly more significant positive results, for example for the momentum anomaly for the US, UK, Germany, and Japan, but results still do not show enough evidence of accepting it as a well working trading strategy. Subsequently looking at the results of the equal weighted averages, we find that the profitability anomaly does show significant positive returns for the Buffett indicator signal, whilst momentum shows a significant positive difference for the CAPE ratio high-low strategy. When looking at all the Sharpe ratios we see almost none of the strategies outperform the market benchmark, except for the Buffett indicator-based cash-based profitability strategy. Turnover ratios for both the Buffett indicator as well as the

Shiller CAPE ratio are good, which means trading costs can be ignored for the most part, if one were to execute the strategies.

5.2) Implications

Due to the above mentioned, implications of this study are limited, since no strong evidence has been found that links the Buffett indicator or the Shiller CAPE ratio with anomaly returns or sentiment proxies. However further research into the predictive power of the CAPE or Buffett indicator can be explored on the foot of this paper. Further research can focus more on the already linked anomalies in this paper for certain countries, where a more-in-depth analysis into the connection between the Buffett indicator and the profitability anomaly or the factor for the Dutch size anomaly can be explored, where research can either focus on explaining why in these countries or why these anomalies specifically seem to connect with the certain Buffett indicator or Shiller CAPE ratio. Research could try and find the determinant other factors that help explain these relationships. Further research can also look to create a different sentiment proxy using both the Buffett indicator and/or CAPE ratio as a part of said sentiment proxy. Since not much is known yet about the predictive power of the Buffett indicator in general, this paper offers a first glance at its predictive power on anomalies. Finally, further research can look into why a cash-based profitability and Buffett indicator strategy outperforms the market, whilst none of the other trading strategies do.

5.3) Limitations

The first limitation of this paper is the lack of sentiment proxies used outside the Baker and Wurgler US sentiment index and US volatility index. Other sentiment proxies for the other countries would further strengthen findings in this paper. After having email contact with the authors of the paper on the world sentiment index, the data was unfortunately not available to be shared. After then trying to compose the sentiment indexes on my own, there was still some data left to be desired and it was unfortunately not possible to accurately create a similar sentiment index for the other countries in the sample. Another limitation could be the exclusion of a prolonged effect of the CAPE ratio of Buffett indicator, which is not included in this paper. Both the Buffett indicator and CAPE are measures of long-term returns and could therefore have a delay in the effect on the anomalies, this could partially help explain the lack of significant results. CAPE data and international anomaly data was also limited, where for France CAPE data was only available since February 1999. The last limitation of this paper is its inability to predict future results, due to the ever-changing market and Buffett indicator values never being as high as now, one cannot with guarantee claim anything about the future returns and predictions in general, even if all returns show significant positive results. Only after a collapse of the market or a fall of the Buffett indicator value, can we truly talk about the predictive power of the Buffett indicator, since such high values may become the new standard.

6) Literature:

Baker, M. and Wurgler, J., (2006) “Investor sentiment and the cross-section of stock returns.” *Journal of Finance* 61, 1645–1680

Baker, M., and Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of economic perspectives*, *21*(2), 129-152.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of financial economics*, *49*(3), 307-343.

Barillas, F., Kan, R., Robotti, C., Shanken, J., (2019). Model comparison with Sharpe ratios. Journal of Financial and Quantitative Analysis, forthcoming, 1–35.

Barroso, P., and Santa-Clara, P. (2015). Momentum has its moments. *Journal of Financial Economics*, *116*(1), 111-120.

BI data other countries: DataStream and https://data-worldbank-org.eur.idm.oclc.org/indicator/NY.GDP.MKTP.CD

Buffett, W. and Loomis, C. (1999) “Mr. Buffett on the Stock Market.”, *Fortune ,* retrieved from : https://archive.fortune.com/magazines/fortune/fortune_archive/1999/11/22/269071/index.htm

Buffett, W., and Loomis, C. (2001) “Mr. Buffett on the Stock Market.”, *Fortune*, retrieved from : https://archive.fortune.com/magazines/fortune/fortune_archive/2001/12/10/314691/index.htm

CAPE ratios: https://indices.barclays/IM/21/en/indices/static/historic-cape.app

Chang, Y. S., and Pak, D. (2018). Warren Buffett value indicator vs. GDP size–is the relationship superlinear? *International Journal of Economics and Business Research*, *15*(2), 223-235.

Dang, T. V., and Xu, Z. (2018). Market sentiment and innovation activities. *Journal of Financial and Quantitative Analysis*, *53*(3), 1135-1161.

Daniel, K. D., Hirshleifer, D. A., and Subrahmanyam, A. (1997). A theory of overconfidence, self-attribution, and security market under-and over-reactions. *Self-Attribution, and Security Market Under-and Over-Reactions (February 19, 1997)*.

Daniel, K., & Hirshleifer, D. (2015). Overconfident investors, predictable returns, and excessive trading. *Journal of Economic Perspectives*, *29*(4), 61-88.

De Bondt, W. F., and Thaler, R. (1985). Does the stock market overreact? *The Journal of finance*, *40*(3), 793-805.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of political Economy*, *98*(4), 703-738.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The journal of Finance*, *25*(2), 383-417.

Fama, E. F., and 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., and French, K. R. (2018). Choosing factors. *Journal of financial economics*, *128*(2), 234-252.

Global anomaly data: https://www.professors.wi.tum.de/fm/research/data/

Hanauer, M. X. (2020). A comparison of global factor models. *Available at SSRN 3546295*.

Henderson, R. (November 7th, 2020) "Berkshire Hathaway rides to $30bn profit on back of investment portfolio. ", *Financial Times,* retrieved from : https://www-ft-com.eur.idm.oclc.org/content/f9cefe0b-759c-4d8e-8ab4-681efcb14f28

Jeffrey Wurgler's webpage: http://people.stern.nyu.edu/jwurgler/

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

Jegadeesh, N., and Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of finance*, *56*(2), 699-720.

Jegadeesh, N., and Titman, S. (2011). Momentum. *Annu. Rev. Financ. Econ.*, *3*(1), 493-509.

Kahneman, D., and Riepe, M. W. (1998). Aspects of investor psychology. *Journal of portfolio management*, *24*(4), 52

Keynes, J. M. (1936). "The general theory of interest, employment and money." *Mcmillan London*

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. Review of Economics and Statistics 47 (February): 13–37

Martin, K. (2021) "Investor anxiety mounts over prospect of stock market 'bubble' " *Financial Times.* Retrieved from: https://www-ft-com.eur.idm.oclc.org/content/a790c796-f0c4-4cf9-8c7a-3b52daff89e4

Miller, E. M. (1977). Risk, uncertainty, and divergence of opinion. *The Journal of finance*, *32*(4), 1151-1168.

Novy-Marx, R., (2010). The other side of value: good growth and the gross profitability premium. Unpublished working paper, University of Chicago'

Rosenberg, B., Reid, K., and Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, *11*(3), 9-16.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance 19 (September): 425–42

Shiller, R. J. (1996). Price-earnings ratios as forecasters of returns: The stock market outlook in 1996. *Unpublished manuscript, Yale University*

Shleifer, A., and Vishny, R. W. (1997). The limits of arbitrage. *The Journal of finance*, *52*(1), 35-55.

Smales, L. A. (2014). News sentiment and the investor fear gauge. *Finance Research Letters*, *11*(2), 122-130.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2012). "The short of it: Investor sentiment and anomalies." *Journal of Financial Economics*, *104*(2), 288-302.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, *104*(2), 288-302.

Stattman, D. (1980). Book values and stock returns. *The Chicago MBA: A journal of selected papers*, *4*(1), 25-45.

Titman, S., Wei, K. J., and Xie, F. (2004). Capital investments and stock returns. *Journal of financial and Quantitative Analysis*, *39*(4), 677-700.

Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, *185*(4157), 1124-1131.

Umlauft, T. S. (2020). The Market Value of Equity-to-Gross Domestic Product Ratio as a Predictor of Long-term Equity Returns: Evidence from the US Market, 1951-2019. *Available at SSRN 3551661*.

US anomaly data: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

US BI data: https://www.longtermtrends.net/market-cap-to-gdp-the-buffett-indicator/

Xing, Y. (2008). Interpreting the value effect through the Q-theory: An empirical investigation. *The Review of Financial Studies*, *21*(4), 1767-1795.