Long-only multi-factor investing in European countries: a comparison of two multi-factor portfolio construction techniques

Roderik Bruins*

May 20, 2023

Abstract

Since around 2010, papers appeared that advise to use the risk factors found in the equity market as building blocks for constructing diversified stock portfolios. Since 2016, several papers investigated and compared two methodologies for constructing multi-factor portfolios. One is called the top-down approach which entails combining single-factor portfolios into a final portfolio. The second is the bottom-up approach and entails selecting stocks that on average have high exposures towards all factors. Several papers concluded that the bottom-up approach is superior compared to the top-down approach on absolute and risk-adjusted basis. Other papers argue there is no differences between both construction techniques. Most papers investigate both construction techniques for the U.S. stock market or for global equities. We contribute to the literature by investigating both construction techniques for fifteen European stock markets over the time period July 2004 till December 2022. Our backtesting results seem to suggest that bottom-up constructed portfolios indeed perform better when applied in a naive manner due to higher factor exposures. However, our backtesting results also seem to suggest that there is no performance difference between both construction techniques when factor exposures are (on average) matched. The backtesting results suggest that both construction techniques can be applied to achieve desired factor exposures.

*SNR: 2065653, e-mail: r.l.w.bruins@tilburguniversity.edu; supervisor: Prof. Dr. B. Melenberg. Tilburg University.

1 Introduction

Markowitz (1952) laid the foundation of *Modern Portfolio Theory* which assumes that investors are risk-averse and therefore want to maximize a portfolio's expected return while minimizing its risk. According to *Modern Portfolio Theory*, one could maximize a portfolio's expected return while minimizing its risk by properly diversifying over multiple assets. Since around 2010, papers appeared in which researchers advise to use risk factors as building blocks for constructing diversified portfolios. These papers argue that risk premiums often have low or negative correlations which could provide diversification benefits. Risk premiums have a cyclic nature and sometimes can even become temporarily negative, but by diversifying over multiple risk factors one could reduce this risk. For example, Ang (2010) is among the first to recognize that risk factors found in the equity market could be used as building blocks for constructing diversified portfolios. Furthermore, Bender et al. (2009), Hjalmarsson (2009), Blitz (2012), Blitz (2015), Brière and Szafarz (2021), and Bessler et al. (2021) backtest multi-factor portfolios and report diversification benefits for multi-factor portfolios.

Bender and Wang (2016) state that the increasingly awareness of using risk factors as building blocks for constructing portfolios raised the question of how to construct such a multi-factor portfolio most efficiently. According to the authors, one could generally distinguish two multi-factor portfolio construction techniques. Bender and Wang (2016) call these construction techniques the top-down approach and the bottom-up approach. Other papers sometimes use other terminologies, for example 'mixing' and 'integrating', or 'portfolio blending' and 'signal blending', but we will follow the terminologies used by Bender and Wang (2016). The top-down approach consists of two steps. In the first step, one creates single-factor portfolios for the risk factors of interest. In the second step, one combines these single-factor portfolios into a multi-factor portfolio by giving each single-factor portfolio a desired weight. The bottom-up approach entails constructing a multi-factor portfolio in just one step by selecting a desired number of stocks that on average have a high exposure towards all risk factors of interest simultaneously.

According to Ghayur et al. (2018), proponents of the bottom-up approach argue that the technique is superior to the top-down approach since it excludes stocks that have favorable exposure to one risk factor but may have unfavorable exposure to one or multiple other risk factors. The risk premium gained by having favorable characteristics to one risk factor therefore may be cancelled out by having unfavorable characteristics to one or multiple other risk factors. Clarke et al. (2016), Bender and Wang (2016), and Fitzgibbons et al. (2017) backtested multi-factor portfolios using both construction techniques. The authors matched portfolios using both construction techniques on a certain risk metric, such as the tracking error or volatility, and find that the bottom-up constructed portfolios outperform in terms of absolute and risk-adjusted return.

Other papers doubt the claims that the bottom-up approach outperform the top-down approach on risk-adjusted basis. Leippold and Rueegg (2018) backtested a lot of top-down and bottom-up constructed portfolios but could not reject the null hypothesis that bottom-up constructed portfolios have statistically significant higher Sharpe Ratios than top-down constructed portfolios when applying a multiple hypothesis framework. The authors therefore argue that the claimed findings favoring the bottom-up approach is a statistical fluke. Ghayur et al. (2018) and Blitz and Vidojevic (2019) argue that papers claiming the superiority of bottom-up constructed portfolios do not make fair comparisons since the tested portfolios are matched on metrics like tracking error or volatility, but do not account for factor exposures. According to the authors, the claimed findings of the superior absolute and risk-adjusted returns for bottom-up constructed portfolios may be simply the results of having higher factor exposures for bottom-up constructed portfolios. The authors conclude that when both types of portfolios are matched on factor exposures, there is no difference between a bottom-up constructed portfolio and a top-down constructed portfolio. Chow et al. (2018) and Amenc et al. (2018) are also skeptical and argue that the findings of superior results of the bottom-up approach are overstated.

The discussion about the superiority of bottom-up constructed portfolios is clearly a recent debate. None of the aforementioned papers test both construction techniques specifically for European stock markets. This paper contributes to the existing literature by backtesting and comparing unleveraged long-only portfolios using both construction techniques for fifteen European stock markets over the time period July 2004 till December 2022. These European stock markets are: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, and United Kingdom. Furthermore, the portfolios will be annually rebalanced at the end of the month June and we make the assumption of zero transaction costs. The multi-factor portfolios will be constructed using the risk factors Size, Value, Profitability, and Investment. Using these four risk factors, it is possible to make a total of eleven combinations of two, three, and four factors. We will backtest for all combinations in order to prevent drawing conclusions on just one or a couple of randomly selected combinations of risk factors.

First, we will backtest portfolios using both construction techniques in a naive setting. We define 'naive setting' as simply selecting a specified percentage (we will also call these percentages 'thresholds') of top ranking stocks for the factors of interest in the investment universe and give an equal weight to each stock. We will do this for thresholds of 10%, 33.33%, and 50%. For clarity, we will call these bottom-up and top-down constructed portfolios the 'naively constructed portfolios' since the portfolios will be constructed by just selecting the top ranking stocks and giving them an equal weight without taking any further considerations into account like matching the portfolio on tracking error, volatility, or factor exposure or like removing unwanted stocks or whatsoever.

By backtesting the naively constructed portfolios, we investigate if bottom-up constructed portfolios perform better than top-down constructed portfolios in a naive setting.

Furthermore, we also backtest average exposures-matched portfolios for both construction techniques in order to investigate whether one portfolio construction technique indeed generates higher risk-adjusted returns than the other. We apply the framework of Ghayur et al. (2018) to construct these average exposures-matched portfolios. The authors construct average exposures-matched portfolios for which they call 'Low level of risk factor exposures' and a 'High level of risk factor exposures'. We will make this distinction as well. The 'Low' levels of average exposures are generated by constructing a bottom-up portfolio that includes the top 50% of the highest ranking stocks in the investment universe. The 'High' levels of average exposures are generated by constructing a bottom-up portfolio that includes 25% of the highest ranking stocks in the investment universe. Consequently, for each bottom-up constructed factor combination portfolio and both levels of exposure, a top-down portfolio is constructed having similar factor exposures for the controlled factors. The threshold of the single-factor portfolios is set such that the average exposures are (roughly) matched over time with the corresponding bottom-up constructed portfolio. For clarity, we will call these portfolios the average exposures-matched portfolios.

Lastly, we also backtest what we call the enhanced constructed portfolios. Blitz and Vidojevic (2019) backtest top-down constructed portfolios having stocks with a specified number of unfavorable characteristics removed from its single-factor portfolios. The authors find that the performance of these portfolios become more similar to bottom-up portfolios when more stocks having negative factor exposures are removed. This implies that a portfolio's performance is the result of factor exposure and not the result of applying a certain portfolio construction technique. We are interested whether we find similar results. Therefore, we will also test naively constructed top-down portfolios having stocks with a specified number of unfavorable characteristics removed. Blitz and Vidojevic (2019) call these 'enhanced strategies'. To make a clear distinction with the naively constructed portfolios and average exposures-matched portfolios, we use this terminology as well and therefore refer to these portfolios as the 'enhanced constructed portfolios'.

The backtesting results show for all thresholds and all factor combinations outperformance of naively constructed bottom-up portfolios compared to naively constructed top-down portfolios in terms of realized factor exposures, annualized mean excess return, and Sharpe Ratio. The results of the backtested naively constructed portfolios therefore seem to be in line with the findings that the bottom-up constructed portfolios outperform top-down constructed portfolios when applied in a naive setting. For the backtested single-factor portfolios we can also see that these portfolios have negative exposure towards at least one factor, in line with the main reason why some favor the bottom-up approach.

For the backtested average exposures-matched portfolios, we find for none of the average exposures-matched bottom-up and top-down constructed portfolios that their difference in Sharpe Ratio and Information Ratio is statistically significant different from zero at 5% significance level. These results seem to contradict the claims that bottom-up constructed portfolios do outperform top-down constructed portfolios on risk-adjusted basis when portfolios controlled for matching factor exposures. Ghayur et al. (2018) and Blitz and Vidojevic (2019) report similar findings. Furthermore, Clarke et al. (2016) match portfolios on tracking error and find higher Information Ratios for bottom-up constructed portfolios than for top-down construction portfolios. We do not actively match on tracking error, however multiple average exposures-matched portfolios have similar tracking errors by coincidence. We find that these portfolios do have similar Information Ratios despite the construction technique applied. And as noted, we could not reject at 5% significance level one of the tested null hypotheses that the difference in Information Ratios are equal to zero. This result also seems to contradict the findings of the authors. Ghayur et al. (2018) note that the ability of top-down constructed portfolios achieving high levels of factor exposures is limited since these portfolios require high concentrations. Initially, we find results supporting this claim. However, the backtested enhanced portfolios show this is not necessarily the case.

Blitz and Vidojevic (2019) find that enhanced top-down constructed portfolios match the performance of bottom-up constructed portfolios. The results of our backtested enhanced top-down constructed portfolios also show a significant performance increase compared to the naively top-down constructed portfolios. Enhanced top-down constructed portfolios only containing stocks with two or more positive factor exposure still lag but approach the performance of naively bottom-up constructed portfolios. Surprisingly, we find that enhanced portfolios containing only stocks having three or four positive factor exposures even have higher factor exposures and consequently outperform the naively bottom-up constructed portfolios.

The backtested results in this paper seem to support the critical views on claims that the bottom-up approach is superior compared to the top-down approach. The backtested results seem to suggest that one could achieve similar risk-adjusted performance when matching on (average) factor exposures. The backtested results also seem to suggest that one could construct top-down portfolios that achieve factor exposures similar to a bottom-up constructed portfolio by adjusting the thresholds of the individual single-factor exposures and/or removing stocks having a specified number of negative factor exposures. The backtested results therefore seem to suggest that one should focus on constructing portfolios having desired factor exposures rather than on construction technique itself.

The remainder of this paper is organized as follows. Section 2 gives an overview of the available and relevant literature. Section 3 gives an overview of the data we use for this research. Section 4 outlines the methodology that we will follow to conduct the research. Section 5 gives an overview of and discusses the found results. Section 6 entails the conclusion.

2 Literature review

The literature review is structured as follows. First, we give an overview of the Modern Portfolio Theory and the research to risk factor models. Secondly, we give an overview of papers advocating the use of risk factors found in the equity market as building blocks for constructing diversified portfolios. Thirdly, we give an overview of two multi-factor portfolio construction techniques. Lastly, we will explain our contribution to the literature.

2.1 Modern Portfolio Theory

Markowitz (1952) laid the foundation of *Modern Portfolio Theory* which assumes that investors are risk-averse and therefore want to maximize a portfolio's expected return while minimizing its risk. According to *Modern Portfolio Theory*, an investor could construct an optimal portfolio by constructing a mean-variance efficient portfolio by properly diversifying over multiple assets. The less, or even better negative, correlation between the invested assets, the better the diversification possibilities.

2.2 Risk factor models

The assumption made by Markowitz (1952) that investors are risk-averse laid the foundation for the *Capital Asset Pricing Model* (CAPM-model) which was independently developed by Sharpe (1964), Lintner (1965), and Mossin (1966). The CAPM-model is the first risk factor model developed for the stock market and explains the cross section of average stock and portfolio returns by only the stock's or portfolio's exposure to the risk factor Market Risk. Risk factor Market Risk is defined as the stocks's or portfolio's sensitivity to the value-weighted market portfolio. In time series this model is expressed as:

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + e_{it}$$

with the null hypothesis that a_i is equal to zero. In this equation, R_{it} is the stock's or portfolio's return over time period t, R_{Ft} is the risk-free rate over time period t, and R_{Mt} is the return of the value-weighted market portfolio over time period t. Furthermore, b_i is the stock's or portfolio's exposure to the risk factor Market Risk, a_i is the return unexplained by the risk factor Market Risk, and e_{it} is the zero-mean residual.

After the introduction of the CAPM-model, multiple papers found that the CAPM-model is not able to explain the cross-sectional expected returns of portfolios that are based on certain firm characteristics. For example, Basu (1977) and Reinganum (1981) are among the first that empirically find that portfolios based on P/E-ratio have expected returns that could not be explained by the CAPM-model. Banz (1981) and Reinganum (1981) are among the first to empirically discover that portfolios based on firm size generate returns that could not be explained by the CAPM-model. Stattman (1980) and Rosenberg et al.

(1985) find for U.S. stocks that the book-to-market ratio have a positive relation with average returns. Bhandari (1988) finds a positive relationship between leverage and average return that could not be explained by the CAPM-model.

Researchers found that extended CAPM-models were able to explain the anomalies, but no model was able to explain all anomalies simultaneously (Hasan et al., 2015). Fama and French (1993) develop the three-factor model by adding the Size factor and the Value factor to the CAPM-model. This model is not only able to explain the identified size and value effects on the cross section of average returns, but also the identified E/P and leverage effects. The authors find that the three-factor model explains over 90 percent of the cross section of average returns of diversified portfolios. The Size factor entails that (a portfolio consisting of) companies with a low market capitalization provide a risk premium over (a portfolio consisting of) companies with a high market capitalization. The Value factor states that (a portfolio consisting of) relatively cheap companies (often defined as companies with a high book-to-market ratio) offers a risk premium over (a portfolio consisting of) relatively expensive companies (often defined as companies with a low book-to-market ratio). In time series this model is expressed as:

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + e_{it}$$

with the null hypothesis that a_i is equal to zero. In this equation, SMB_t is the return of a diversified portfolio consisting of stocks with low market capitalization minus the return of a diversified portfolio consisting of stocks with high market capitalization over time period t and HML_t is the return of a diversified portfolio consisting of stocks with a high book-to-market value minus the return of a diversified portfolio consisting of stocks with a low book-to-market values over time period t. Next to that, s_i and h_i are the stock's or portfolio's exposures to the risk factors Size and Value respectively. Furthermore, a_i is the return unexplained by the risk factors Market Risk, Size, and Value.

Fama and French (1993) state that the three-factor model fails to explain much of the variation in average returns linked to the profitability of companies and their investments. Haugen and Baker (1996) and Cohen et al. (2002) find, when controlling for book-to-market value, a positive relation between the profitability of firms and average returns. Fairfield et al. (2002), Richardson and Sloan (2003), and Titman et al. (2004) find a negative relation between invested capital and average stock returns. Fama and French (2006) derive from the dividend discounting model, developed by Miller and Modigliani (1961), the following equation:

$$\frac{M_t}{B_t} = \frac{\sum_{\tau=1}^{\infty} E[Y_{t+\tau} - dB_{t+\tau}]/(1+r)^{\tau}}{B_t}$$

In this equation, M_t denotes the market capitalization at time t, B_t is the total book equity at time t, $Y_{t+\tau}$ is the total equity earnings for time period $t+\tau$, $dB_{t+\tau} = B_{t+\tau} - B_{t+\tau-1}$ denotes the change in total book equity for time period

$t + \tau$, and r denotes the stock's expected return over a time period τ. This equation suggests that a stock's expected return is related to the company's current book-to-market value, expected profitability, and expected investments. To be more specific, the equation suggests a positive relationship between a company's total equity earnings and the company's expected stock return, a negative relationship between a company's total book equity growth and the company's expected stock return, and a positive relationship between a company's book-to-market ratio and the company's expected stock return. The authors observe the suggested relation but fail to empirically find a significant negative relationship between stock returns and expected investment. Aharoni et al. (2013) argue that the failure in the empirical work of Fama and French (2006) is due to using per-share level variables instead of firm level variables. The authors claim per-share level variables are not reliable due to share issuance or repurchase and therefore firm level variables should be used. The authors get similar empirical findings of Fama and French (2006) when using per-share level variables, but do find the significant negative relation between stock returns and expected investment when using firm level variables. Novy-Marx (2013) criticizes the usage of current earnings as proxy for future profitability as used by Fama and French (2006). Novy-Marx (2013) states that current earnings can give a false representation of a firm's profitability. Highly profitably firms could have low current earnings since investments (like research and development, human capital development, and advertising) are treated as expenses resulting in low(er) current earnings. The author therefore argues that the gross profits is the best accounting measure for profitability. The author uses the gross profits-to-assets ratio as proxy for future profitability and finds it has strong predictive power for the average stock returns.

Using the findings of Aharoni et al. (2013) and Novy-Marx (2013), Fama and French (2015) propose a five-factor model by adding the risk factors Profitability and Investment to their three-factor model. The resulting five-factor model therefore consists of the five risk factors Market Risk, Size, Value, Profitability, and Investment. The Profitability factor entails that (a portfolio consisting of) companies with robust profitability outperforms (a portfolio consisting of) companies with a weak profitability. The authors define companies with robust profitability as companies having high values for operating profitability minus interest expense divided by its book equity, and subsequently companies with weak profitability as companies having low values for the operating profitability minus interest expense divided by its book equity. The Investment factor entails that (a portfolio consisting of) companies with conservative investment policy outperforms (a portfolio consisting of) companies with an aggressive investment policy. A company with conservative investment policy barely invests (often measured as low total asset growth) while a company with an aggressive investment policy invests a lot (often measured as high total asset growth). In time series this model is expressed as:

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - R_{Ft}) + s_i SMB_t + h_i HML_t + r_i RMW_t + c_i CMA_t + e_{it}$$

with the null hypothesis that a_i is equal to zero. In this equation, RMW_t is the return of a diversified portfolio consisting of companies with robust profitability minus the return of a diversified portfolio consisting of companies with a weak profitability over time period t. CMA_t is the return of a diversified portfolio consisting of companies with a conservative investment policy minus the return of a diversified portfolio consisting of companies with an aggressive investment policy over time period t. Furthermore, r_i and c_i are the stock's or portfolio's exposure to the risk factors Profitability and Investment respectively. Furthermore, a_i is the return unexplained by the risk factors Market Risk, Size, Value, Profitability, and Investment.

2.3 Risk factors in equity market as diversification building blocks

In line with *Modern Portfolio Theory*, it is common practice to diversify across multiple (uncorrelated) assets, asset classes, company sectors and geographical regions in order to reduce a portfolio's risk. However, since around 2010 papers appeared in which it was recognized that the risk factors could also be used as building blocks for constructing diversified investment portfolios. Ghayur et al. (2018) state that risk premiums have a cyclic nature and sometimes can even become temporarily negative. The authors also state that risk factors often have low or negative correlations which could provide diversification benefits. Ang (2010) is among the first to recognize that the risk factors could also be used to diversify. Bender et al. (2009) compare over the time period May 1995 till September 2009 a multi-factor portfolio with a portfolio having 60 percent invested in equities and 40 percent in bonds. For both portfolios the authors use global ETF's. The authors find for both portfolios similar annualized excess returns but an almost 3.5 times lower annualized volatility for the multi-factor portfolio. The authors therefore suggest that combining risk factors could be more attractive than traditional asset allocations. Asness et al. (2013) find across eight diverse markets and asset classes consistent risk premiums for risk factors value and momentum. They also find a negative correlation between the Value and Momentum factors, which holds across all markets and asset classes. Due to the negative correlation, the authors recognize its potential for constructing a diversified portfolio using the Value and Momentum factors. The authors find for all considered markets and asset-classes for the time period 1970's till July 2011 that a diversified long-short strategy having an equal weight on the Value and Momentum factors provided substantial diversification benefits. Hjalmarsson (2009) extends this approach by also considering the following characteristics: short-term reversals, long-term reversals, cashflow-price ratio, earnings-price ratio, and size. The author finds for the U.S. market over the time period 1951 till 2008 that a characteristic-based strategy having an equal weight on each of the seven single-characteristic long-short portfolios almost always had a higher Sharpe ratio than all single-characteristic long-short portfolio strategies or the long-short momentum-value characteristic-based portfolio as noted by Asness et al. (2013). Blitz (2012) also advises to use risk factors as building block to diversify a portfolio. The author tests for the U.S. market over the time period July 1963 until December 2009 a long-only portfolio that is equally weighted in the risk factors Market Risk, Value, Momentum, and Low-Volatility and find that it improved the performance compared to a market portfolio. Ilmanen and Kizer (2012) also advocate using factor diversification over asset-class diversification. The authors compare a factor diversification strategy with an asset-class diversification strategy over the time period 1973-2010. The authors find that a long-short multi-factor portfolio invested in the U.S. market and diversified over factors Value, Momentum, Carry, and Trend achieved a Sharpe ratio of 1.44, while a portfolio that is diversified over U.S. and global asset-classes only achieved a Sharpe ratio of 0.48. Blitz (2015) repeats the strategy tested by Blitz (2012) for the U.S. market and finds that it also outperformed the market over the time period 2010-2014. Blitz (2015) also tests multi-factor portfolios using the risk factors Profitability and Investment proposed by Fama and French (2015). The author tests multi-factor portfolios that are equally weighted in the risk factors Value, Momentum, Low-Volatility, Operating Profitability, and Investment for U.S. stocks over the time period July 1963 till December 2014 but do not find performance improvement or deterioration. Brière and Szafarz (2021) compare multi-factor portfolios with passive sector investing for the U.S. stock market over the time period 1963 till 2014. The multi-factor portfolios use the risk factors Size, Value, Profitability, Investment, and Momentum as building blocks. The authors find that diversification over risk factors, when short-selling is unrestricted, outperforms sector diversification. Without short-selling, factor diversification only performs better than sector diversification during bull periods but underperforms during bear periods. Bessler et al. (2021) extend the research of Brière and Szafarz (2021) by applying various portfolio optimization approaches on the factor diversified portfolios and the sector diversified portfolios. Bessler et al. (2021) use ETF's to backtest the portfolios in the U.S. stock market. For the time period May 2007 till November 2020, the authors find that for longer investment horizons the factor portfolios delivered better performances. They also find that during 'normal' time the factor diversified portfolios clearly outperforms the sector diversified portfolios, but during bear periods the sector diversified portfolios offers better diversification.

2.4 Construction techniques for multi-factor portfolios

Due to increased popularity of multi-factor investing, the question raised how to construct such a multi-factor portfolio in a most efficient way (Bender & Wang, 2016). Bender and Wang (2016) distinguish two multi-factor portfolio construction techniques which they call the top-down approach and the bottom-up approach. The top-down approach is done by first constructing multiple single-factor portfolios and then combining these single-factor portfolios into one portfolio. The bottom-up approach entails that the investor creates a portfolio consisting of a certain number of individual stocks that on average have a high exposure to the risk factors of interest. Several other names are used for both techniques in other papers. The top-down approach is for example in other papers also called the mixing approach or the portfolio blending approach. The bottom-up approach is for example in other papers also called the integrating approach or the signal blending approach. In this paper, we will use the names given by Bender and Wang (2016).

Ghayur et al. (2018) state that proponents of the bottom-up approach argue that it achieves better factor exposure than the top-down approach. This because each single-factor portfolio includes stocks that have favorable characteristics towards the risk factor of interest, but ignores the fact that they may also have unfavorable characteristics towards other risk factors. For such stocks, the risk premium gained by having favorable characteristics on one risk factor may be cancelled out by having unfavorable characteristics towards one or multiple other risk factors. Proponents of the top-down approach however argue that the strategy is more transparent and offers better insight into the performance contributions.

Several papers investigate whether one method is better than the other. Some papers claim that the bottom-up approach deliver superior results compared to top-down constructed portfolios. Bender and Wang (2016) test both multi-factor portfolio construction techniques using the risk factors Value, Momentum, Quality, and Low Volatility over the time period 1993 till March 2015. The authors find that the bottom-up constructed portfolio had a 20 percent higher Information Ratio than the top-down constructed portfolio. Clarke et al. (2016) test both construction techniques for 1,000 U.S. stocks over the time period 1968 till 2015 using the risk factors Low Beta, Size, Value, and Momentum. The authors find that a long-only bottom-up constructed portfolio achieved about 20 percent higher Sharpe Ratio and Information Ratio compared to a long-only top-down constructed portfolio. Fitzgibbons et al. (2017) test long-only portfolios using both construction techniques for a universe of stocks that roughly corresponds to the MSCI World Index-benchmark over the time period February 1993 till December 2015 using the risk factors Value and Momentum. The authors find for all levels of tracking error that bottom-up constructed portfolios have higher Information Ratios than top-down portfolios when both types of portfolios are matched on tracking error. Only for very low levels of tracking error (below 1%), the authors find similar performance. Lester (2019) tests portfolios using a wide range of factor definitions and finds that the bottom-up approach increasingly outperform the top-down approach when a large number of low correlated factors are used. The author finds that a bottom-up portfolio using four orthogonal factors produces twice the factor exposure and outperformance compared to a top-down portfolio and 40 percent higher Information Ratio.

Other papers are skeptical about the claims that the bottom-up approach is superior. Leippold and Rueegg (2018) argue that claims of bottom-up constructed portfolios achieving higher risk-adjusted returns compard to top-down constructed portfolios are violating the standard finance theories in which one can only achieve higher returns by taking more risk. The authors are therefore skeptical about these claims. The authors test long-only portfolios using both construction techniques for all NYSE, AMEX, and NASDAQ stocks over the time period June 1963 till December 2016. The portfolios are constructed by using 26 combinations of the risk factors Value, Profitability, Investment, Momentum, and Low Volatility. They find that bottom-up approach indeed had better performance than the top-down approach. However, when applying a multiple hypothesis framework they could not reject the null hypothesis that bottom-up constructed portfolios generate statistically significant higher Sharpe Ratios than top-down constructed portfolios. Chow et al. (2018) test both construction techniques using Value, Momentum, Profitability, Investment, and Low-Beta. The authors find that the bottom-up approach outperformed the top-down approach in terms of absolute and risk-adjusted returns. They find however that the bottom-up approach has a higher concentration in stocks with the consequence of having higher idiosyncratic active risk and higher estimated transaction costs than the top-down approach. They argue that the top-down approach is more transparent, flexible and offers better insight into the performance contributions. Therefore, the authors recommend using the top-down approach. Amenc et al. (2018) argue that, due to possible over-fitting and multiple testing biases, the backtests of bottom-up approach may be overstated. Next to that, they argue that the bottom-up approach results in too high concentrated portfolios. Therefore, the authors argue that one should be critical about the claims that the bottom-up approach is superior. Ghayur et al. (2018) and Blitz and Vidojevic (2019) argue that the papers claiming the superiority of the bottom-up approach do not make fair comparisons. The authors state that these papers compare both construction techniques by matching them on a risk metric such as tracking error or volatility. However, according to the authors it would only be a fair comparison when the portfolios are matched on factor exposures. The authors argue that the claimed findings of the superior absolute and risk-adjusted returns for bottom-up constructed portfolios may be simply the results of having higher factor exposures for bottom-up constructed portfolios. According to Blitz and Vidojevic (2019), claiming that the bottom-up approach is better on risk-adjusted basis compared to the top-down approach would suggest that there is some sort of 'bottom-up premium' (the authors call it 'integration premium' since they use other terminology for the construction techniques) i.e. one could earn higher risk-adjusted returns by applying the bottom-up approach without taking more risk. Ghayur et al. (2018) compare long-only top-down and bottom-up constructed portfolios that are matched on average factor exposures. They use the Russell 1000 Index universe and the Value, Momentum, Quality, and Volatility risk factors. The authors find for high levels of factor exposures, that bottom-up constructed portfolios have higher Information Ratios compared to top-down constructed portfolios while both being average exposures-matched. For low-to-moderate levels of factor exposure, they find that top-down constructed portfolios often have higher Information Ratios compared to bottom-up constructed portfolios while both being average exposures-matched. However, they also note that these differences show little significance. The authors state that when factor exposures are (roughly) matched, the superiority of the bottom-up approach could be challenged. Blitz and Vidojevic (2019) test both portfolio construction approaches and an in-between approach using stocks traded on NYSE, AMEX, and NASDAQ exchanges for the time period 1963 till 2017. They consider the risk factors Size, Value, Profitability, Investment, and Momentum. The in-between approach is equal to the top-down approach but has stocks with a specified number of unfavorable characteristics towards other risk factors removed from the single-factor portfolios. The authors find that the top-down approach is suboptimal in obtaining factor exposures. However, the authors argue that it is to be expected since bottom-up constructed portfolios are more concentrated and therefore more effective in capturing higher factor exposures and consequently outperform top-down constructed portfolios. The authors argue that when portfolios are matched on factor exposures, the applied construction technique does not matter. They demonstrate this by using the in between approach. They find that with this approach the portfolios could achieve similar performance compared to bottom-up constructed portfolios.

2.5 Factor allocation strategies

Asness (2016) notes that it would be very attractive to 'time' the factors of interest such that one can give the factors with the highest conditional expected returns the highest weights in the portfolio and vice versa. However, the author comes to the conclusion that it is very difficult to predict the premiums of risk factors and therefore very hard to time factors. The author argues that trying to time factors is like trying to time the market which is very hard, if not impossible, to do and should therefore be avoided. Therefore, he advises to stick only to factors that are backed by scientific evidence (and particularly out-of-sample evidence) and economic theory and of which the investor believes will persist on the long run. Asness et al. (2017) also conclude that successfully timing factor exposure is even harder than successfully timing asset class exposure, while the latter is already hard to do. Dichtl et al. (2019) find that it is possible to time factors using fundamental and technical time-series predictors, but since it results in a high turnover it is very hard to profit from it after transaction costs. Dichtl et al. (2020) investigate various factor-based allocation strategies for constructing multi-factor portfolios using global equities. The authors evaluate 17 different strategies over the period from January 2006 to December 2019. They conclude that the naive equally-weighted strategies cannot be outperformed by more sophisticated allocation strategies.

2.6 Contribution to literature

Several papers compare the top-down and bottom-up construction techniques for multi-factor portfolios. Clarke et al. (2016), Leippold and Rueegg (2018), Ghayur et al. (2018), and Blitz and Vidojevic (2019) test both construction techniques for the U.S. market. Fitzgibbons et al. (2017) and test both techniques for global equities.

The discussion about the superiority of bottom-up constructed portfolios is clearly a recent debate. Clarke et al. (2016), Bender and Wang (2016), and Fitzgibbons et al. (2017) empirically find that the bottom-up approach outperforms the top-down approach on absolute and risk-adjusted basis. Also Lester (2019) finds that bottom-up constructed portfolios outperform top-down constructed portfolios. Amenc et al. (2018) and Chow et al. (2018) do also find superior performance for the bottom-up approach but are skeptical and argue that the claimed superiority of bottom-up constructed portfolios is overstated. Leippold and Rueegg (2018) find using a multiple hypothesis framework that bottom-up constructed portfolios do not generate statistically significant higher Sharpe Ratios than top-down constructed portfolios and conclude that the superiority of bottom-up constructed portfolios is a statistical fluke. Ghayur et al. (2018) and Blitz and Vidojevic (2019) argue that the papers claiming the superiority of the bottom-up approach do not make fair comparisons since they do not compare (average) exposure matched portfolios. After comparing (average) exposure-matched portfolios, the authors find that both types of portfolios become very similar in performance.

The inconsistent findings of the aforementioned papers show that it is clearly a debate whether bottom-up constructed portfolios are superior compared to top-down constructed portfolios or not. Furthermore, none of the aforementioned papers test both construction techniques specifically for European stock markets. This paper contributes to the existing literature by investigating for fifteen European stock markets over the time period July 2004 till December 2022 whether unleveraged long-only multi-factor portfolios constructed using the bottom-up approach could have achieved higher absolute and risk-adjusted returns compared to multi-factor portfolios constructed using the top-down approach.

3 Data

As data source for all stock data and fundamental data Thomson Reuters Refinitiv Datastream is used. Thomson Reuters Refinitiv Datastream is an extensive database containing lots of global economic data including global equity data and global company fundamentals. Stock data and fundamental data of the European public companies is extracted from the Datastream database by using the Worldscope lists provided by Datastream. These lists contain per country all companies included in their database that are or were publicly traded. Therefore, also public companies that stopped existing or are delisted are included. Therefore, the obtained data is survivorship-bias free.

Fama and French use sixteen European stock markets in order to construct their research portfolios with for the European stock market. The research conducted in this paper will be applied on almost the same stock markets. We use the stock markets of the following countries: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, and United Kingdom. Furthermore, the Datastream database is also used to extract the monthly shareholder returns of the *DAX*-index, *SP Europe 350*-index, *STOXX 600*-index. Next to that, the Datastream database is also used to extract the returns of German government bonds with a maturity of 1 month and the returns of European government bonds with a maturity of 1 month. The monthly shareholder returns of the Fama French European market portfolio are downloaded from the website of Kenneth French[1].

For the time period July 2004 till December 2022, the monthly shareholder returns (with reinvested dividends) and the factor scores are gathered from Datastream for all public companies in the selected countries. Next to that, for each month it is determined whether the company has a stock price in order to determine whether a company was active in that month. The factor scores are calculated on an annual basis. The factor scores are calculated in the same fashion as Fama and French (2015) do. For each company, the Size factor score for year t is calculated as the market value at the end of the month June of year t. The Value factor score is determined by the book-to-market ratio using the book value for the fiscal year ending in year $t-1$ and the market value at the end of the month December of year $t-1$. The Profitability factor score for year t is calculated by dividing the annual revenues minus cost of goods sold, interest expense, and selling, general, and administrative expenses by the book equity using the accounting data for the fiscal year ending in year $t-1$. The Investment factor score for year t is calculated as the growth of total asset for the fiscal year ending in $t-1$ divided by the total assets for the fiscal year ending in $t-2$. For all non-Euro countries, the monthly shareholder returns are converted to returns in terms of Euros and the Size scores are converted to Euros.

[1]https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data$_{l}$*ibrary.html*

All public companies having missing values for the stock price, monthly shareholder returns, or calculated factor scores for the entire time period January 2002 till December 2022 are filtered out. This is for example the case for companies that stopped being publicly traded before January 2002 or for public companies for which the required accounting data to calculate the factor scores with is not available for the entire time period January 2002 till December 2022. Public companies that existed in the time period January 2002 till December 2022 but stopped being publicly traded before December 2022 are not filtered out in order to prevent the survivorship-bias.

Table 1 presents for each stock market the number of companies available in our dataset before and after filtering. Table 2 presents the number of stocks available in our stock universe at each moment of rebalancing.

TABLE 1: Number of companies per stock market available in our dataset before and after filtering

Stock market	**Number of companies before filtering**	**Number of companies after filtering**
Austria	228	76
Belgium	362	98
Denmark	653	198
Finland	347	168
France	2044	396
Germany	1947	842
Greece	470	258
Ireland	189	77
Italy	883	460
Netherlands	502	199
Norway	762	294
Portugal	163	14
Spain	540	134
Sweden	765	715
United Kingdom	6016	2135
Total	**15.871**	**6.064**

This table shows per country the number of stocks available in our dataset. The column 'Number of companies before filtering' denotes the number of companies available in our dataset before filtering out companies with missing values for the entire time period January 2002 till December 2022 for monthly shareholder returns, stock price, or the factor scores. The column 'Number of companies after filtering' contain the number of companies in our dataset after filtering. These are the companies being used to backtest with.

TABLE 2: Number of companies available in the stock universe at each moment of rebalancing.

Moment of rebalancing	**Number of stocks in universe**
End of June 2004	631 stocks
End of June 2005	633 stocks
End of June 2006	692 stocks
End of June 2007	762 stocks
End of June 2008	796 stocks
End of June 2009	794 stocks
End of June 2010	775 stocks
End of June 2011	782 stocks
End of June 2012	757 stocks
End of June 2013	735 stocks
End of June 2014	763 stocks
End of June 2015	771 stocks
End of June 2016	797 stocks
End of June 2017	793 stocks
End of June 2018	787 stocks
End of June 2019	800 stocks
End of June 2020	811 stocks
End of June 2021	813 stocks
End of June 2022	878 stocks
Average	**767 stocks**

This table shows the number of stocks in the universe available at each moment of rebalancing.

4 Methodology

In this section, the methodology of the research used in this paper is outlined. This section is structured as follows. First, a general overview of the backtesting is given. Secondly, the procedure how we construct the multi-factor portfolios is outlined. Thirdly, we make a selection for the benchmark. Fourthly, we will give an overview of the metrics we use and how we calculate them. Fifthly, we outline which metrics we will not use and explain why. In the last part, we give an overview of the (types of) portfolios we are going to backtest.

4.1 Backtesting

The research in this paper will be done by means of backtesting. A backtesting program written in R will be used to backtest the long-only portfolios over the time period July 2004 till December 2022. The portfolios will be annually rebalanced at the end of the month June of year t starting in year 2004. The rebalancing moments are chosen such that they correspond to the rebalancing moments of the Fama and French research portfolios. Although we could also have chosen other rebalancing moments, or multiple rebalancing per year for example, we found this most intuitive. Since the factor scores Value, Profitability, and Investment are calculated using accounting data of the fiscal year ending in year $t-1$, this creates a lag of a minimum of 6 months in order to make sure that the information used is publicly available at moment of rebalancing.

The time period July 2004 till December 2022 is selected in order to analyse recent data but also to make sure that all factor scores are measured in Euros and not in a mix of Euros and local currencies. All countries in the Eurozone officially switched to the Euro currency on 1st of January 2002. However, the Investment factor for example is measured as the relative total asset growth between year $t-1$ and $t-2$. In order to measure the total asset growth in terms of Euros and not in terms of a mix of Euros and a local currency, the starting moment of backtesting is chosen to be the last trading day of June 2004. Furthermore, we only consider unleveraged long-only portfolios and we make the assumption of zero transaction costs.

4.2 Construction of the long-only multi-factor portfolios

In this paper, the portfolio construction techniques as for example applied in Leippold and Rueegg (2018) and Ghayur et al. (2018) will be used. Each stock obtains for each risk factor a score which is based on its corresponding factor score. Please note that 'factor score' and 'score' are different here. According to Leippold and Rueegg (2018), there are two common methodologies for calculating such a score which they call the rank-based approach and the z-score approach. The rank-based approach entails using the security's rank within the investment universe for the factor of interest in order to derive a score. The z-score approach involves transforming a security's factor score into a z-score.

We will make use of the z-score approach since it will also be used to measure the factor exposures. This will be further outlined in section 4.4. For year t, a long-only portfolio using k number of risk factors as building blocks will be rebalanced according to the following procedure:

1. In the first step, all companies that do not have a stock price for December of year $t-1$ or June of year t are filtered out. Next to that, companies with missing values for year $t-1$ for total assets, revenues, costs of goods sold, general and administrative expenses or interest expense are left out. Furthermore, companies with negative book equity for year $t-1$ and missing value for total assets for year $t-2$ are left out. $\Phi \in \mathbb{R}^{n \times k}$ contains for all filtered securities $i = 1, ..., n$ the factor scores for the risk factors of interest $f = 1, ..., k$. Each column $\phi_f \in \mathbb{R}^n$ of matrix Φ contains for all n securities the factor score corresponding to factor f.

2. In the second step, for each stock i, the z-score Z_i^f is calculated for each considered factor f using the factor scores of year t. The z-scores are calculated by applying the following formula:

$$Z_i^f = \frac{\phi_{f,i} - \mu(\phi_f)}{\sigma(\phi_f)}$$

   In this equation, function $\mu(\phi_f)$ calculates the mean of column ϕ_f and function $\sigma(\phi_f)$ calculates the standard deviation of column ϕ_f.

   The market capitalizations and positive book-to-market ratios are heavily right skewed and take the form of a log-normal distribution. Therefore, we use the natural logs of these factor scores to calculate the z-scores with as also done by for example Blitz and Vidojevic (2019). Furthermore, at each moment of rebalancing there are companies with extreme unfavorable values for both the Profitability and Investment factor (companies with extreme low profitability or companies with extreme high asset growth) causing the distributions of the profitability and investment factor scores also being heavily skewed. These outliers therefore result in undesired distorted z-scores for the Profitability and Investment factor since they cause a high $\sigma(\phi_f)$ resulting in low z-scores. Since these companies are very unattractive to invest in, we filter out companies with extreme profitability and investment scores using IQR filtering. This results in desirable z-score distributions.

3. In the third step the rebalanced portfolio is constructed. Let $w_{td}, w_{bu} \in \mathbb{R}^n$ contain the stocks weights for all n stocks in a top-down constructed portfolio and bottom-up constructed portfolio respectively. Furthermore, $Z^f \in \mathbb{R}^n$ is a vector containing the z-scores of all securities in the investment universe for factor f and $\varphi(.)$ is a function that uses the (composite) z-score to calculate the stock weights of the single- or multi-factor portfolio. According to Leippold and Rueegg (2018), weights a_f and function

$\varphi(.)$ could be freely chosen. For long-only and unleveraged top-down constructed portfolios however, the values for a_f and the sum of a_f should be greater than or equal to 0 and less than or equal to 1. For bottom-up constructed portfolios, a_f is just a factor exposure's scalar and therefore any value could be chosen in order to construct portfolios with desired factor exposures. For the backtested portfolios in this paper, we will specify function $\varphi(.)$ in subsection 4.5. As noted in the literature review, multiple papers advise to apply a naive equally weighted strategy regarding the factor weights (see for example also Asness (2016), Asness et al. (2017), Dichtl et al. (2019) and Dichtl et al. (2020)). In this paper, this advise is followed by giving for both construction techniques an equal weight to each risk factor i.e. $a_f = \frac{1}{k}$. Depending on the construction technique applied, the portfolio will be constructed as follows:

(a) **Top-down construction technique:**
This construction technique involves a two-step procedure:

1. First, for each factor of interest a single-factor portfolio is created having stock weights $w_f \in \mathbb{R}^n$:

$$w_f = \varphi\left(Z^f\right)$$

2. Secondly, all single-factor portfolios are combined into one portfolio by giving each single-factor portfolio a weight a_f:

$$w_{td} = \sum_{f=1}^{k} a_f w_f$$

(b) **Bottom-up construction technique:**
This construction technique involves a one-step procedure in which each factor is given a weight a_f in order to derive for each stock a composite z-score which will be used to construct a portfolio:

$$w_{bu} = \varphi\left(\sum_{f=1}^{k} a_f Z^f\right)$$

4.3 Benchmark

The German *DAX*-index is chosen as benchmark to compare the performance of the backtested multi-factor portfolios with. Several indices are considered to serve as benchmark. The benchmark should be a fair representation of the European stock market. Therefore, the German *DAX*-index, the *SP Europe 350*-index, the *SP Europe 350*-index, and the Fama French European market portfolio are considered as potential benchmarks. Since the performance of the backtested (multi-)factor portfolios will be compared to the performance of an index that could be passively followed, an index that has a high annualized return and annualized Sharpe ratio (using European government bonds with maturity of 1 month) over the time period July 2004 till December 2022 has to be be chosen. In table these statistics are summarized for the considered benchmarks.

	Annualized mean excess return	Annualized Sharpe Ratio
DAX	7.34%	0.45%
SP Europe 350	6.48%	0.47%
STOXX 600	6.61%	0.47%
FF market portfolio	5.63%	0.41%

TABLE 3: Comparison of European stock indices over the time period July 2004 till December 2022.

The table shows that the *SP Europe 350*-index and the *STOXX 600*-index have the highest annualized Sharpe ratio. The *DAX*-index however has by far the highest annualized return and just a slightly lower annualized Sharpe ratio than the *SP Europe 350*-index and *STOXX 600*-index. Since the *DAX*-index has the highest annualized return and not a too poor annualized Sharpe ratio compared to the other indices, *DAX*-index is chosen to serve as benchmark.

Since the *DAX*-index is chosen as benchmark, it would be a good choice to use German government bonds with a maturity of one month as proxy for the European monthly risk-free rate. However, over the time period June 2004 till October 2010 there is no data available for these bonds. Over the time period November 2010 till December 2022 the correlation between the German government bonds with a maturity of one month and European government bonds with maturity of one month was 99.9%. Due to this high correlation, we choose to use European government bonds with maturity of one month as proxy for the German bonds with maturity of one month and therefore as proxy for the risk-free rate.

4.4 Used metrics

We are mainly interested in comparing the performance of portfolios in terms of absolute returns and risk-adjusted returns. Therefore, the key performance statistics in this paper are the *ex-post* annualized mean excess return, annualized excess volatility, and the Sharpe Ratio. Furthermore, we assess portfolios by considering the average number of companies and the realized average factor exposures. We also utilize the active return, active risk, and Information Ratio but in lesser extent. Please note that we write the metrics in population form but in fact we calculate the metrics using the backtested sample data.

4.4.1 Annualized mean excess return

In this paper, we define the excess return as a portfolio's return in excess to the risk-free rate. Since this paper investigates the performance of multi-factor portfolios relative to the benchmark, the mean of the active return is an important metric to assess whether a strategy on average outperforms a benchmark in terms of absolute returns. The mean active return is defined as:

$$\text{Mean excess return} = E\left[R_p - R_f\right]$$

In this equation, R_p denotes the annualized returns of a backtested portfolio and R_f denotes the annualized risk-free rates.

4.4.2 Annualized excess risk

We define excess risk as the volatility of a portfolio's excess return. It is a measure for the amount of risk taken by a portfolio. The excess risk is defined as:

$$\text{Excess risk} = \sigma\left(R_p - R_f\right)$$

In this equation, R_p denotes the returns of a backtested portfolio and R_f denotes the risk-free rates.

4.4.3 Annualized Sharpe Ratio

The Sharpe ratio is a widely used metric to analyse the past performance of a portfolio. It is a measure of how much return a portfolio has generated in excess to the risk free rate per unit of risk taken in excess to the risk free rate. Sharpe (1994) defines the ex post Sharpe ratio as follows:

$$\text{Sharpe ratio} = \frac{\text{Annualized mean excess return}}{\text{Annualized excess risk}} = \frac{E\left[R_p - R_f\right]}{\sigma\left(R_p - R_f\right)}$$

The annualized Sharpe Ratio will be used in this paper because it enables to compare the risk-adjusted returns of the backtested portfolios compared to the risk-free rate.

4.4.4 Annualized mean active return

The active return is a portfolio's return in excess to a benchmark's return. The mean of the active return is an metric used to assess whether a strategy on average outperforms a benchmark in terms of absolute returns. The mean active return is defined as:

$$\text{Active return} = E\left[R_p - R_b\right]$$

In this equation, R_p denotes the returns of the backtested portfolios and R_b denotes the returns of the benchmark.

4.4.5 Annualized active risk

Active risk, also known as the tracking error, is the standard deviation of the returns of a portfolio in excess to the benchmark returns. The active risk therefore is a measure for riskiness of a portfolio compared to a benchmark. The active risk is defined as:

$$\text{Active return} = \sigma\left(R_p - R_b\right)$$

In this equation, R_p denotes the returns of the backtested portfolios and R_b denotes the returns of the benchmark.

4.4.6 Annualized Information Ratio

The Information Ratio is a measure for much return a portfolio has generated in excess to a benchmark per unit of excess risk taken compared to the benchmark. The Information Ratio is calculated by dividing the mean active return over the active risk:

$$\text{Information ratio} = \frac{\text{Active return}}{\text{Active risk}} = \frac{E\left[R_p - R_b\right]}{\sigma\left(R_p - R_b\right)}$$

In this equation, R_p denotes the returns of the backtested portfolios and R_b denotes the returns of the benchmark.

4.4.7 Mean factor exposures

Multi-factor portfolios aim to have above-average exposure towards the risk factors of interest. Therefore, it is necessary to know the (mean) factor exposures of a portfolio in order to get a sense of how successful a portfolio is in capturing factor exposures. According to Ghayur et al. (2018) a portfolio's exposure towards a factor can be measured by the weighted z-scores or by the active risk contribution. The authors state that both methodologies can also be used to construct a multi-factor portfolio. According to them, the first method is a more direct and easier to implement approach and therefore they use the first method. This methodology is therefore also used in this paper. Ghayur et al. (2018) calculate the exposure of a portfolio to a risk factor f at time t, consisting of n stocks with each stock i having a weight w_i and having a z-score Z_i^f for risk factor f, as follows:

$$\text{Portfolio's exposure at time } t \text{ to risk factor } f = \sum_{i=1}^{n} w_i \mathrm{Z}_i^f$$

The portfolio's mean exposure towards a risk factor is calculated by taking the average exposure to the risk factor at all T moments of rebalancing:

$$\text{Portfolio's average exposure to risk factor } f = \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} w_{t,i} \mathrm{Z}_{t,i}^f$$

4.4.8 Average number of companies in portfolio

In this paper, the average number of companies in a portfolio is defined as the average number of companies included in the portfolio at time of rebalancing. In this paper, this metrics is not especially used to analyse a portfolio's performance but merely to give a sense of how diversified a backtested portfolio is in comparison to other backtested portfolios.

However, it could give a sense of the amount of risk taken. The more (uncorrelated) companies included in a portfolio, the less risky a portfolio is in general. The risk of a portfolio could be dissected into the exposure to companies' idiosyncratic risks and exposure to systematic risks. When more (uncorrelated) companies are included into a portfolio, the portfolio's exposure to the idiosyncratic risk of individual companies is reduced and therefore the overall portfolio's risk is reduced. Therefore, assuming the companies are uncorrelated the number of companies included in the backtested portfolio gives a sense of the riskiness of the portfolio.

Next to that, it gives a sense of a portfolio's ability to capture factor exposures. Ghayur et al. (2018) and Blitz and Vidojevic (2019) state that higher factor exposures could only be more effectively achieved when having more concentrated portfolios. Therefore, the mean number of companies in a portfolio also gives a sense of the portfolio's ability to achieve higher factor exposures.

4.5 Unused metrics

Other often used metrics to assess a portfolio's performance are for example the arithmic mean of returns, geometric mean of returns, volatility, Treynor Ratio, skewness of returns, kurtosis of returns, and Value at Risk (VaR) of returns. These metrics will not be used in this paper.

The arithmic and geometric mean of returns do say something about a portfolio's absolute return and a portfolio's volatility does say something about a portfolio's absolute risk. Although these metrics could be used to compare portfolio performances, we will utilize the mean excess return and excess risk since we also use the Sharpe Ratio. Therefore, it does not add a lot of value by also using the arithmic mean, geometric mean, and volatility as well.

The Treynor Ratio is defined as a portfolio's active risk divided by its exposure towards the risk factor Market Risk. Although it could give a sense of how much factor exposure a portfolio has beyond it's exposure to Market Risk, it does not seem to add much value for this paper since it is not a precise measure for how much exposure a portfolio has towards other factors than the Market Risk. Since we will also measure the portfolio's explicit factor exposures, it is not necessary to include this ratio.

Furthermore, the metrics skewness, kurtosis and VaR are generally used to assess a portfolio's risk of suffering exceptional (negative) returns. In this paper, we are only interested in comparing portfolio's exposure towards risk factors, the mean excess return, and the mean risk-adjusted return.

4.6 Portfolios to be backtested

4.6.1 Considered factors

This paper follows the advice given by Asness (2016) to stick only to a few risk factors of which the investor believes will generate persistent and robust returns on the long term when constructing multi-factor portfolios. Cochrane (2011) notes that hundreds of potential risk factors have been reported in the literature over the last decades. However, a lot of these claimed findings are likely the result of data mining and therefore be false (Harvey et al., 2016). Therefore, in this paper only factors that has been broadly acknowledged by the academic literature will be used. To that end, the following factors will be used: Size, Value, Profitability and Investment. The risk factors Momentum, identified by Jegadeesh and Titman (1993), and Low-Volatility, identified by Haugen and Heins (1972), were considered as well but we decided to not include these factors.

To prevent drawing conclusions based on findings that are simply the result of combining a certain combination of factors, this paper will backtest portfolios constructed using all possible combinations of 2, 3 and 4 risk factors using the risk factors Size ('S'), Value ('V'), Profitability ('P') and Investment ('I'). Using 2 factors, it is possible to construct 6 combinations. For 3 factors it is possible to construct 4 combinations and for 4 factors it is possible to construct 1 portfolio. Therefore, it is possible to construct in total 11 possible factor combinations. This leads to the following set of possible factor combinations that will be backtested:

$$\mathcal{C} = \{SV, SP, SI, VP, VI, PI, SVP, SVI, SPI, VPI, SVPI\}$$

To clarify, a portfolio using the factor combination 'SV' only controls for the risk factor Size and Value. The factor Profitability and Investment are left uncontrolled.

4.6.2 Naively constructed multi-factor portfolios

First, non-factor exposure matched multi-factor portfolios will be backtested in order to get a sense of how both construction techniques perform when applied in a naive setting. Blitz and Vidojevic (2019) test portfolios based on selecting the top 10%, 20%, 33.33%, or 50% (we will call these numbers also 'thresholds') highest ranking stocks for the factors of interest in the investment universe. In line with their methodology, in this paper three portfolios will be backtested that invest in the top 50% ('50% portfolio'), the top tercile ('33.33% portfolio'), and the top decile ('10% portfolio') ranking stocks of the investment universe in order to be able to investigate the effect across several thresholds. These thresholds are chosen such that the differences in thresholds are approximately as large as possible. The number of testing just three thresholds is arbitrarily chosen. Furthermore, testing for three different thresholds serves as a robustness check in order to prevent drawing conclusions based on the results of just one threshold.

After selected the stocks, these will subsequently be given an equal weight. According to Blitz and Vidojevic (2019) it is also common practice to give the selected stocks a value weight. However, the main implication is that it gives a negative tilt towards the Size factor and therefore reducing returns. This has also been found in backtests we performed for this paper. Therefore, only equally-weighted portfolios will be considered.

For clarity, we will call these portfolios the 'naively constructed portfolios' since the portfolios will be constructed by just selecting the top ranking stocks and giving them an equal weight without taking any further considerations into account like matching the portfolio on tracking error, volatility, or factor exposure or like removing unwanted stocks or whatsoever.

To this end, for both strategies $\varphi(.)$ is set such that the top 50%, 33.33%, and 10% ranking stocks are given an equal-weight and the other stocks in the investment universe are given a weight of zero. Since we will backtest portfolios for 2 portfolio construction techniques, 3 thresholds and for 11 factor combinations, a total of 66 naively constructed multi-factor portfolios will be backtested.

4.6.3 Average exposures-matched portfolios

According to Ghayur et al. (2018) and Blitz and Vidojevic (2019), a fair comparison between the top-down approach and the bottom-up approach could only be made when the factor exposures are (roughly) similar to each other. To that end, in this paper also average exposures-matched portfolios will be backtested. To construct average exposures-matched portfolios, the framework proposed by Ghayur et al. (2018) will be used. The authors suggest that one could construct average exposures-matched top-down portfolios by varying the thresholds of the single-factor portfolios used to construct the top-down portfolio. The authors argue that one could achieve average exposures-matched portfolios by either matching the exposures at time of rebalancing or by matching the average exposures over time. The authors use the latter approach, which will also be used in this paper. The authors also tested portfolios matching at time of rebalancing but found the portfolios had similar performance.

For each factor combination, we construct bottom-up portfolios and top-down portfolios that (roughly) have the same mean factor exposures over time for the factors used to build up the portfolio. Ghayur et al. (2018) construct average exposures-matched portfolios for which they call 'Low level of risk factor exposures' and a 'High level of risk factor exposures'. In this paper, this distinction will be made as well. The 'Low' levels of average exposures are generated by constructing a bottom-up portfolio that includes the top 50% of the stocks. The 'High' levels of average exposures are generated by constructing a bottom-up portfolio that includes 25% of the stocks. For both average exposure levels, a top-down portfolio using the same factor combinations is constructed. The single-factor portfolios will have equal weight in the top-down portfolio. The average exposures of the top-down constructed portfolios are matched with the average exposures of the bottom-up portfolio by varying the thresholds for each of the considered single-factor portfolios individually.

It is important to note that the average factor exposures of the portfolios are only controlled for the considered factor combinations. The exposures to the other factors remain uncontrolled. For example, for SV-portfolios are only constructed using the risk factor Size and Value while (the exposures to) the risk factors Profitability and Investment remain uncontrolled. The VPI-portfolios for example are controlled for the risk factors Value, Profitability, and Investment while leaving (the exposures to) uncontrolled.

We will compare the Sharpe Ratios and Information Ratios to determine whether both construction techniques generate similar risk-adjusted returns when average-exposure matched. For each pair of average exposures-matched portfolios, we will test the null hypothesis that the difference between the two Sharpe Ratios is equal to zero. Furthermore, this will also be done for the Information Ratios. To be more specific, we will test the following hypothesis:

$$H_0 : \Delta = 0 \text{ vs } H_1 : \Delta \neq 0$$

In this equation, Δ denotes the difference between the true Sharpe Ratios of two portfolios. For the hypothesis test regarding the Information Ratios, Δ denotes the difference between the true Information Ratios. To test this null hypothesis, we will use the method proposed by Ledoit and Wolf (2007). Leippold and Rueegg (2018) and Ghayur et al. (2018) also use this methodology. Ledoit and Wolf (2007) argue that Sharpe Ratio testing methodologies proposed by Jobson and Korkie (1981) and Memmel (2003) are not appropriate in case the returns are not normally distributed or are serial correlated. According to Ledoit and Wolf (2007), the former is often the case in financial returns and therefore argue that these methodologies are not valid. The authors therefore argue that it is better to use a heteroskedasticity and autocorrelation consistent (HAC) covariance matrix estimator proposed by Andrews (1991) and Andrews and Monahan (1992). However, the authors also state that the inference is less accurate in case of having small to moderate sample sizes. They note that many papers demonstrated that the inference accuracy could be greatly improved when using studentized bootstrap. Therefore, the authors propose a methodology based on studentized circular block bootstrapping to generate a two-sided confidence interval using M-number of bootstrap samples to test the null hypothesis that the difference between the Sharpe Ratios of two portfolios is equal to zero.

The authors also provide an algorithm to determine the optimal block size to generate the bootstrap samples with. Since we have a finite sample size (222 monthly returns for each average exposures-matched portfolio) we decide to use the algorithm outlined by Ledoit and Wolf (2007) to test whether average exposures-matched portfolios perform similarly in terms of Sharpe Ratios and Information Ratios. As noted by the authors, this method could in the same manner be applied for Information Ratios.

To test the null hypothesis and subsequently calculate the p-values, we will use the R-function *sharpeTesting* from the R-package *PeerPerformance*[2] as also used by Ardia and Boudt (2018). We will apply the settings such that it corresponds to the methodology as outlined in Ledoit and Wolf (2007). Furthermore, for each hypothesis test we use 5000 circular block bootstrap samples to generate the confidence intervals with. This number corresponds to the number of circular block bootstrap samples as used in an example given by Ledoit and Wolf (2007).

4.6.4 Enhanced constructed multi-factor portfolios

Lastly, we are interested in what will happen to the performance of naively constructed top-down portfolios when one would remove from the utilized single-factor portfolios stocks having a specified number of negative factor exposures. Blitz and Vidojevic (2019) do this as well and find that it significantly improves the returns of a top-down constructed portfolio. To be more specific, the authors find that when stocks with more negative exposures are removed from the single-factor portfolios, the more the factor exposures and consequently the performance of the top-down constructed portfolios become similar to that of a bottom-up constructed portfolio. The authors therefore argue that there is no such thing as a 'bottom-up premium'. In addition to the average exposures-matched portfolios, we are interested in whether we can find this in our dataset as well. Blitz and Vidojevic (2019) call the top-down portfolios having stocks with a specified number of unfavorable characteristics removed 'enhanced strategies'. To make a clear distinction with the naively constructed and average exposures-matched portfolios, we will use this terminology as well. Therefore, we call these portfolios the 'enhanced constructed portfolios' since these are naively constructed portfolio but having stocks with a certain number of unfavorable characteristics removed and therefore being a bit more advanced.

[2]The R-codes of this package could be found on https://CRAN.R-project.org/package=PeerPerformance

5 Results

In this section, the results of all backtested portfolios will be discussed. First, the results of the backtested naive multi-factor portfolios will be discussed. In the second part, the results of the backtested average exposure matched multi-factor portfolios will be discussed. In the third part, the results of the enhanced multi-factor portfolios will be discussed.

5.1 Naively constructed multi-factor portfolios

Tables 4 and 5 show the results for the backtested naively constructed portfolios. The tables show that for all thresholds, the average backtested naively constructed bottom-up portfolio does have a higher mean excess return and Sharpe Ratio compared to the average backtested naively constructed top-down portfolio. The backtesting results suggest that bottom-up constructed portfolios are superior to the top-down constructed portfolios when constructed in a naive manner. For the average backtested 50% portfolios, the differences in the performance statistics are not very large for both construction techniques. To be more specific, the average backtested naively constructed bottom-up portfolio with a threshold of 50% has a mean excess return of 9.87% and a Sharpe Ratio of 0.63, while the average backtested naively constructed top-down portfolio with a threshold of 50% has a mean excess return of 9.31% and a Sharpe Ratio of 0.60. However, when the thresholds are tightened the performance differences between the average backtested naive top-down constructed portfolio and the average backtested naive bottom-up constructed portfolio increase in favor for the latter. To be more specific, for the 33.33% threshold, the average bottom-up portfolio has a mean excess return of 10.88% and a Sharpe Ratio 0.67 and the average top-down portfolio has a mean excess return of 9.71% and a Sharpe Ratio of 0.62. For the 10% threshold, the average bottom-up portfolio has a mean excess return of 13.87% and a Sharpe Ratio 0.71 and the average top-down portfolio has a mean excess return of 11.54% and a Sharpe Ratio of 0.65.

Blitz and Vidojevic (2019) and Leippold and Rueegg (2018) also find that the bottom-up constructed portfolio outperform top-down constructed portfolios when applied in a naive manner. According to Blitz and Vidojevic (2019), these findings are to be expected since for the same factor combinations a naively constructed bottom-up portfolio contain much fewer companies than a naively constructed top-down portfolio and are therefore much more concentrated. According to the authors, higher concentrated portfolios are better able to capture high factor exposures. The results in tables 4 and 5 show that this is indeed the case and are therefore in line with the authors' view. The average naively bottom-up constructed portfolio contains on average 297, 199, and 60 companies for thresholds 50%, 33.33%, and 10% respectively. The average naively top-down constructed portfolio however contains on average 488, 378, and 135 companies for thresholds 50%, 33.33%, and 10% respectively. According to

Ghayur et al. (2018) and Blitz and Vidojevic (2019), the fact that bottom-up constructed portfolios perform better is therefore a direct consequence of the fact that they simply have higher factor exposures.

Table 4: Backtested results of naively constructed bottom-up portfolios

Naively constructed bottom-up equally-weighted top 50% portfolios								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	9.20	15.90	0.58	297.40	0.61	0.58	-0.29	0.14
SP	10.50	15.40	0.68	297.40	0.43	0.04	0.36	-0.02
SI	8.30	15.70	0.52	297.40	0.57	0.30	-0.25	0.51
VP	11.50	15.90	0.72	297.40	0.09	0.40	0.35	0.02
VI	9.30	15.90	0.58	297.40	0.28	0.56	-0.27	0.52
PI	9.70	15.30	0.64	297.40	-0.15	-0.13	0.45	0.41
SVP	11.00	15.90	0.69	297.40	0.51	0.47	0.11	0.06
SVI	8.60	15.80	0.54	297.40	0.53	0.53	-0.30	0.45
SPI	10.00	15.40	0.65	297.40	0.43	0.13	0.17	0.44
VPI	10.70	15.60	0.68	297.40	0.13	0.40	0.18	0.45
SVPI	9.80	15.70	0.62	297.40	0.48	0.47	0.01	0.40
Avg	**9.87**	**15.68**	**0.63**	**297.40**	**0.36**	**0.34**	**0.05**	**0.31**

Naively constructed bottom-up equally-weighted top 33.33% portfolios								
Factors combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	10.50	17.20	0.61	198.50	0.79	0.82	-0.39	0.23
SP	11.90	15.50	0.77	198.50	0.54	0.03	0.56	-0.04
SI	8.50	16.30	0.52	198.50	0.78	0.39	-0.39	0.70
VP	13.00	16.50	0.79	198.50	0.18	0.52	0.49	0.03
VI	10.10	16.70	0.60	198.50	0.41	0.77	-0.38	0.68
PI	10.60	15.60	0.68	198.50	-0.20	-0.22	0.69	0.50
SVP	12.10	16.60	0.73	198.50	0.69	0.65	0.15	0.10
SVI	9.90	16.50	0.60	198.50	0.73	0.72	-0.43	0.61
SPI	10.20	15.50	0.66	198.50	0.56	0.15	0.28	0.56
VPI	11.70	16.30	0.72	198.50	0.21	0.54	0.24	0.59
SVPI	11.20	16.50	0.68	198.50	0.65	0.64	0.01	0.55
Avg	**10.88**	**16.29**	**0.67**	**198.50**	**0.49**	**0.46**	**0.08**	**0.41**

Naively constructed bottom-up equally-weighted top 10% portfolios								
Factors combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	15.20	23.80	0.64	59.80	1.17	1.35	-0.61	0.36
SP	13.20	17.30	0.76	59.80	0.65	-0.11	1.22	-0.09
SI	10.50	18.90	0.56	59.80	1.12	0.44	-0.67	1.20
VP	16.40	17.40	0.95	59.80	0.33	0.71	0.92	0.06
VI	11.40	21.50	0.53	59.80	0.68	1.20	-0.59	1.06
PI	10.80	15.70	0.69	59.80	-0.17	-0.34	1.23	0.75
SVP	19.50	20.60	0.95	59.80	0.99	1.00	0.38	0.17
SVI	13.30	22.30	0.60	59.80	1.10	1.18	-0.65	0.92
SPI	11.60	17.70	0.66	59.80	0.74	0.08	0.71	0.85
VPI	15.10	19.70	0.76	59.80	0.38	0.76	0.53	0.88
SVPI	15.60	21.00	0.74	59.80	0.96	0.99	0.14	0.81
Avg	**13.87**	**19.63**	**0.71**	**59.80**	**0.72**	**0.66**	**0.24**	**0.63**

This table shows the results of backtested naively bottom-up constructed portfolios. The upper table shows the results for the portfolios using a threshold of 50%, the middle table shows the results for the portfolios using a threshold of 33.33%, and the lower table shows the results for the portfolios using a threshold of 10%. Each table's last row contain for each column the average value in bold letters. The column 'Factor combinations' denotes the specific factors used to construct a portfolio with while leaving the other factors uncontrolled. The used risk factors are Size ('S'), Value ('V'), Profitability ('P'), and Investment ('I'). For example, a portfolio using factor combination 'VI' uses risk factors Value and Investment while a portfolio using the factor combination 'SVP' use the risk factors Size, Value, and Profitability. The column 'Mean excess return' denotes a portfolio's annualized mean return in excess to the risk-free rate. The column 'Excess risk' denotes the annualized volatility of a portfolio's returns in excess to the risk-free rate. The column 'Sharpe Ratio' contains for each portfolio the Sharpe Ratio which is calculated by dividing a portfolio's mean excess return by its excess risk. The fifth column denotes the average number of companies a portfolio holds. The sixth, seventh, eighth, and ninth column denote a portfolio's average factor exposure towards the risk factors Size, Value, Profitability, and Investment respectively.

Table 5: Backtested results of naively constructed top-down portfolios

Naively constructed top-down equally-weighted 50% single-factor portfolios								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	9.10	15.90	0.57	402.40	0.50	0.50	-0.26	0.13
SP	9.60	15.10	0.63	484.50	0.22	0.00	0.22	0.00
SI	8.20	15.50	0.53	424.80	0.43	0.23	-0.20	0.40
VP	10.40	15.40	0.68	493.90	0.00	0.20	0.21	0.01
VI	9.00	15.70	0.58	421.20	0.21	0.43	-0.21	0.41
PI	9.50	15.20	0.62	467.00	-0.08	-0.07	0.27	0.27
SVP	9.70	15.40	0.63	546.30	0.24	0.24	0.06	0.05
SVI	8.80	15.70	0.56	476.40	0.38	0.39	-0.22	0.31
SPI	9.10	15.20	0.60	535.40	0.19	0.05	0.10	0.22
VPI	9.70	15.40	0.63	544.40	0.04	0.19	0.09	0.23
SVPI	9.30	15.40	0.61	566.20	0.21	0.22	0.00	0.20
Avg	**9.31**	**15.45**	**0.60**	**487.50**	**0.21**	**0.22**	**0.01**	**0.20**

Naively constructed top-down equally-weighted 33.33% single-factor portfolios								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	9.90	16.70	0.59	295.40	0.70	0.66	-0.36	0.18
SP	9.80	15.00	0.65	357.80	0.31	-0.04	0.32	0.00
SI	8.20	15.80	0.52	309.70	0.59	0.27	-0.29	0.53
VP	11.20	15.60	0.72	369.80	0.05	0.26	0.30	0.02
VI	9.70	16.30	0.59	310.40	0.33	0.57	-0.30	0.55
PI	9.50	15.00	0.63	347.00	-0.06	-0.13	0.38	0.37
SVP	10.30	15.60	0.66	439.00	0.35	0.29	0.08	0.07
SVI	9.20	16.20	0.57	370.60	0.54	0.50	-0.32	0.42
SPI	9.20	15.10	0.61	432.40	0.28	0.03	0.14	0.30
VPI	10.10	15.50	0.65	444.50	0.10	0.23	0.12	0.31
SVPI	9.70	15.50	0.63	483.50	0.32	0.26	0.01	0.27
Avg	**9.71**	**15.66**	**0.62**	**378.19**	**0.32**	**0.26**	**0.01**	**0.27**

Naively constructed top-down equally-weighted 10% single-factor portfolios								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	13.20	21.20	0.62	98.90	1.09	1.05	-0.59	0.33
SP	11.60	16.00	0.72	117.10	0.54	-0.04	0.54	0.07
SI	10.50	18.40	0.57	108.10	0.91	0.36	-0.55	0.90
VP	12.50	17.60	0.71	117.30	0.22	0.42	0.55	0.07
VI	11.50	19.40	0.59	109.20	0.59	0.83	-0.54	0.90
PI	9.80	15.50	0.63	115.90	0.04	-0.26	0.59	0.64
SVP	12.50	17.70	0.70	155.70	0.62	0.48	0.17	0.16
SVI	11.80	19.30	0.61	141.30	0.86	0.74	-0.56	0.71
SPI	10.60	16.00	0.66	162.50	0.50	0.02	0.19	0.54
VPI	11.30	17.00	0.67	164.90	0.28	0.33	0.20	0.53
SVPI	11.60	17.30	0.67	195.20	0.57	0.39	0.00	0.48
Avg	**11.54**	**17.76**	**0.65**	**135.10**	**0.57**	**0.39**	**0.00**	**0.48**

This table shows the results of backtested naively top-down constructed portfolios. The upper table shows the results for top-down constructed portfolios using single-factor portfolios with a threshold of 50%, the middle table shows the results for the top-down constructed portfolios using single-factor portfolios with a threshold of 33.33%, and the lower table shows the results for top-down constructed portfolios using single-factor portfolios with a threshold of 10%. Each table's last row contain for each column the average value in bold letters. The column 'Factor combinations' denotes the specific factors used to construct a portfolio with while leaving the other factors uncontrolled. The used risk factors are Size ('S'), Value ('V'), Profitability ('P'), and Investment ('I'). For example, a portfolio using factor combination 'VI' uses risk factors Value and Investment while a portfolio using the factor combination 'SVP' use the risk factors Size, Value, and Profitability. The column 'Mean excess return' denotes a portfolio's annualized mean return in excess to the risk-free rate. The column 'Excess risk' denotes the annualized volatility of a portfolio's returns in excess to the risk-free rate. The column 'Sharpe Ratio' contains for each portfolio the Sharpe Ratio which is calculated by dividing a portfolio's mean excess return by its excess risk. The fifth column denotes the average number of companies a portfolio holds. The sixth, seventh, eighth, and ninth column denote a portfolio's average factor exposure towards the risk factors Size, Value, Profitability, and Investment respectively.

5.2 Average exposures-matched portfolios

Table 6 presents the results for the backtested average exposures-matched portfolios and table 8 presents for each pair of backtested average exposures-matched bottom-up and top-down constructed portfolios the p-values for the hypothesis test with the null hypothesis that the difference between the Sharpe Ratios of both portfolios is equal to zero. As mentioned in section 4, the exposures are only average-matched for the controlled factors. For example, the SI portfolios are only average exposures-matched for the Size and Investment factors, and not for the Value and Profitability factors. Although the portfolios are only matched on the risk factors of the specified risk factor combination, for several backtested portfolios the exposures towards the uncontrolled risk factors are roughly similar as well for both construction techniques.

An interesting finding is that for none of the 22 pairs backtested factor combinations and factor exposure levels the null hypothesis is rejected at 5% significance level. However, two out of the 22 null hypotheses are rejected at 10% significance level. The first is the Value-Profitability portfolios at high level of factor exposures having a p-value of 0.050. The second is the Size-Investment portfolios at high level of factor exposures having a p-value of 0.070. However, when applying the multiple hypothesis framework proposed by Holm (1979), none of the null hypotheses are rejected at 10% significance level. According to this framework, the lowest p-value should be below 0.10/22=0.0045 which is not the case. In other words, it could not be rejected that the portfolios perform significantly different to each other on risk-adjusted basis. These backtested results seem to be in line with findings of recent papers that are skeptical about claims that the bottom-up approach is superior to the top-down approach on risk-adjusted basis. Blitz and Vidojevic (2019) find that portfolios constructed with similar factor exposures, despite being constructed using the bottom-up approach or the top-down approach, have similar performance. The authors find portfolios having similar factor exposures achieve similar returns and risk-adjusted returns despite the portfolio construction technique applied. Leippold and Rueegg (2018) find that, even for non (average) exposures-matched portfolios, there are no statistical significant differences between the Sharpe Ratios of bottom-up constructed portfolios and top-down constructed portfolios that use the same factor combinations when applying a multiple hypothesis framework. These results contrasts the findings of Clarke et al. (2016), Bender and Wang (2016), and Fitzgibbons et al. (2017) that bottom-up constructed portfolios do perform better on absolute and on risk-adjusted basis.

Although none of the backtested average exposures-matched pairs are rejected at 5% significance level, the table shows that the mean excess returns and excess risks do differ for both construction techniques. This is most likely explained by the fact that we match the average factor exposures over time and not at time of rebalancing like Blitz and Vidojevic (2019) do. Blitz and Vidojevic (2019) argue that portfolios, despite the construction technique applied, having similar factor exposures do have very similar performances. Ghayur et al. (2018) also construct their portfolios average exposures-matched. However, the authors note they also tested also portfolios being factor exposure matched at time of rebalancing and found similar results. Therefore, it is to be expected that when the factor exposures are matched at time of rebalancing the excess return, excess risk, and Sharpe Ratio of both types of portfolios will become very similar as well. Therefore, I expect that the same conclusions will hold for portfolios being factor exposure-matched at time of rebalancing instead of being average exposures-matched.

In table 7, the Information Ratios are displayed for the backtested average exposures-matched portfolios. Table 8 presents for each pair of backtested average exposures-matched portfolios the p-values for the hypothesis test with the null hypothesis that the difference between the Information Ratios is equal to zero. The table shows that none of the average exposures-matched portfolio pairs have Information Ratios that are statistically significant different to each other at both 5% and 10% significance level. This contradicts the findings of Ghayur et al. (2018). The authors find for two, three and four factor combinations, that for low-to-moderate levels of factor exposures the top-down approach generates higher Information Ratios than the bottom-up approach, while for high levels of factor exposures the bottom-up constructed portfolios generate higher Information Ratios. The results of our backtested average exposures-matched portfolios are not in line with their findings. Although the authors also state that the found differences in Information Ratios show little statistical significance.

The findings of Ghayur et al. (2018) contrasts the finding of Fitzgibbons et al. (2017). The authors of the latter paper find that bottom-up constructed portfolios do have higher Information Ratios for all levels of tracking errors. Only for very low values of tracking error the authors find the bottom-up and top-down portfolio perform similarly. It is important to note that in this paper the backtested average exposures-matched portfolios are not actively matched on tracking error. However, table 7 shows that by accident the bottom-up and top-down constructed portfolios for factor combinations SPI (low), SV (low), SP (low), SI (low), VP (low), and PI (low) do in fact have (roughly) similar tracking errors for both construction techniques. The table also shows that these portfolios have very similar Information Ratios for both construction techniques. In matter of fact, none of the differences in Information Ratio are statistically significant different from zero. Also interesting, the at low exposure level top-down constructed portfolio using the factor combination SP has similar tracking error as the at low exposure level bottom-up constructed portfolio using the factor combination PI. However, the former portfolio realized an Information Ratio of 0.11 while the latter portfolio realized an Information Ratio of 0.06. These findings also contrast the findings of Fitzgibbons et al. (2017) that bottom-up portfolios with similar tracking errors to top-down portfolios perform better. The authors only use the Value and Momentum factors to construct their portfolios. According to Blitz and Vidojevic (2019), it may be possible that these performance differences are simply the results of unknowingly having unintended differences in exposures towards uncontrolled risk factors.

Furthermore, table 6 shows that for both construction techniques similar levels of factor exposures could be achieved. This is in line with the findings of Ghayur et al. (2018). However, the table also shows that for several factor combinations the top-down constructed portfolios require much higher concentrated (single-factor) portfolios in order to achieve the same factor exposures as the bottom-up constructed portfolios. This is most often the case for high levels of average factor exposures. This is the case for seven out of the eleven bactested average exposures-matched portfolio pairs. To be more specific, this is the case for the factor combinations SVPI, SVP, SPI, VPI, SP, VP, and PI. Although similar average factor exposures could be achieved for both portfolio construction techniques, this would suggest that there is a limit for top-down portfolios in achieving similar factor exposures to bottom-up constructed portfolios. This is in line with Ghayur et al. (2018) who notes that, although the factor exposures could be matched using both construction techniques, the ability for top-down constructed portfolios to achieve (very) high levels of factor exposures is limited since it would require high concentrations in the single-factor portfolios. These findings therefore would suggest that bottom-up constructed portfolios potentially could achieve higher factor exposures. However, the results of the backtested enhanced portfolios will show this is not necessarily the case.

Table 6: Backtested results of average exposures-matched portfolios and their Sharpe Ratios

Factor combination	Exposure level	Construction technique	Ann. mean excess return	Ann. excess risk	Sharpe Ratio	Avg. no. companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SVPI	Low	Top-down	10.93	16.99	0.64	242	0.52	0.42	0.01	0.44
SVPI	Low	Bottom-up	9.80	15.75	0.62	297	0.48	0.47	0.01	0.40
SVPI	High	Top-down	12.66	19.53	0.65	129	0.66	0.55	-0.04	0.54
SVPI	High	Bottom-up	11.07	16.34	0.68	208	0.63	0.61	0.01	0.53
SVP	Low	Top-down	12.02	17.65	0.68	269	0.52	0.48	0.10	0.12
SVP	Low	Bottom-up	10.97	15.87	0.69	297	0.51	0.47	0.11	0.06
SVP	High	Top-down	16.01	23.10	0.69	104	0.72	0.68	0.17	0.15
SVP	High	Bottom-up	12.65	16.99	0.74	179	0.73	0.69	0.16	0.11
SVI	Low	Top-down	9.15	16.77	0.55	330	0.57	0.56	-0.37	0.54
SVI	Low	Bottom-up	8.62	15.84	0.54	297	0.53	0.53	-0.30	0.45
SVI	High	Top-down	11.10	18.72	0.59	172	0.81	0.72	-0.52	0.65
SVI	High	Bottom-up	10.08	16.83	0.60	179	0.77	0.78	-0.46	0.64
SPI	Low	Top-down	10.02	15.37	0.65	264	0.43	0.03	0.13	0.44
SPI	Low	Bottom-up	9.98	15.40	0.65	297	0.43	0.13	0.17	0.44
SPI	High	Top-down	15.08	20.89	0.72	79	0.65	0.02	0.28	0.61
SPI	High	Bottom-up	11.11	15.41	0.72	149	0.62	0.16	0.37	0.65
VPI	Low	Top-down	11.56	17.34	0.67	220	0.27	0.39	0.15	0.46
VPI	Low	Bottom-up	10.65	15.64	0.68	297	0.13	0.40	0.18	0.45
VPI	High	Top-down	12.65	19.95	0.63	94	0.38	0.53	0.24	0.54
VPI	High	Bottom-up	11.70	16.30	0.72	208	0.20	0.52	0.23	0.57
SV	Low	Top-down	9.57	16.26	0.59	341	0.62	0.59	-0.32	0.16
SV	Low	Bottom-up	9.16	15.89	0.58	297	0.61	0.58	-0.29	0.14
SV	High	Top-down	11.38	19.72	0.58	173	0.93	0.96	-0.50	0.27
SV	High	Bottom-up	10.81	18.31	0.59	149	0.91	0.96	-0.46	0.27
SP	Low	Top-down	10.97	15.19	0.72	254	0.42	-0.04	0.35	0.03
SP	Low	Bottom-up	10.50	15.39	0.68	297	0.43	0.04	0.36	-0.02
SP	High	Top-down	12.56	15.92	0.79	88	0.55	-0.07	0.68	0.06
SP	High	Bottom-up	12.05	15.45	0.78	149	0.57	-0.01	0.72	-0.07
SI	Low	Top-down	8.17	15.77	0.52	323	0.57	0.26	-0.27	0.51
SI	Low	Bottom-up	8.26	15.75	0.52	297	0.57	0.30	-0.25	0.51
SI	High	Top-down	10.09	17.61	0.57	133	0.86	0.34	-0.51	0.82
SI	High	Bottom-up	7.75	16.50	0.47	149	0.88	0.43	-0.47	0.83
VP	Low	Top-down	11.72	16.35	0.72	258	0.13	0.35	0.35	0.04
VP	Low	Bottom-up	11.50	15.91	0.72	297	0.09	0.40	0.35	0.02
VP	High	Top-down	12.64	18.39	0.69	74	0.26	0.53	0.60	0.07
VP	High	Bottom-up	13.80	16.34	0.84	149	0.21	0.58	0.59	0.04
VI	Low	Top-down	9.51	16.47	0.58	285	0.35	0.60	-0.33	0.58
VI	Low	Bottom-up	9.27	15.91	0.58	297	0.28	0.56	-0.27	0.52
VI	High	Top-down	11.27	20.02	0.56	123	0.58	0.88	-0.49	0.81
VI	High	Bottom-up	10.89	17.71	0.61	149	0.50	0.89	-0.44	0.79
PI	Low	Top-down	9.66	14.98	0.64	273	-0.04	-0.16	0.43	0.43
PI	Low	Bottom-up	9.74	15.30	0.64	297	-0.15	-0.13	0.45	0.41
PI	High	Top-down	11.09	15.74	0.70	89	0.07	-0.29	0.85	0.61
PI	High	Bottom-up	10.10	15.38	0.66	149	-0.21	-0.28	0.83	0.56

This table shows backtest results for the average exposures-matched portfolios. The column 'Factor combinations' denotes the specific factors used to construct a portfolio with while leaving the other factors uncontrolled. The used risk factors are Size ('S'), Value ('V'), Profitability ('P'), and Investment ('I'). For example, a portfolio using factor combination 'VI' uses risk factors Value and Investment while a portfolio using the factor combination 'SVP' use the risk factors Size, Value, and Profitability. The second column denotes whether the matched exposure are low or high. The third column denotes the construction technique applied. The column 'Ann. mean excess return' denotes a portfolio's annualized mean return in excess to the risk-free rate. The column 'Ann. excess risk' denotes the annualized volatility of a portfolio's returns in excess to the risk-free rate. The column 'Sharpe Ratio' contains for each portfolio the Sharpe Ratio which is calculated by dividing a portfolio's annualized mean excess return by its annualized excess risk. The seventh column denotes the average number of companies a portfolio holds. The eighth, ninth, tenth, and eleventh column denote a portfolio's average factor exposure towards the risk factors Size, Value, Profitability, and Investment respectively.

Table 7: Backtested results of average exposures-matched portfolios and their Information Ratios

Factor combination	Exposure level	Construction technique	Ann. active return	Ann. active risk	Information Ratio	Avg. no. companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SVPI	Low	Top-down	2.47	24.22	0.10	242	0.52	0.42	0.01	0.44
SVPI	Low	Bottom-up	1.34	23.38	0.06	297	0.48	0.47	0.01	0.40
SVPI	High	Top-down	4.19	25.90	0.16	129	0.66	0.55	-0.04	0.54
SVPI	High	Bottom-up	2.60	23.94	0.11	208	0.63	0.61	0.01	0.53
SVP	Low	Top-down	3.55	24.70	0.14	269	0.52	0.48	0.10	0.12
SVP	Low	Bottom-up	2.51	23.38	0.11	297	0.51	0.47	0.11	0.06
SVP	High	Top-down	7.54	28.46	0.26	104	0.72	0.68	0.17	0.15
SVP	High	Bottom-up	4.18	24.21	0.17	179	0.73	0.69	0.16	0.11
SVI	Low	Top-down	0.68	24.11	0.03	330	0.57	0.56	-0.37	0.54
SVI	Low	Bottom-up	0.16	23.46	0.01	297	0.53	0.53	-0.30	0.45
SVI	High	Top-down	2.64	25.66	0.10	172	0.81	0.72	-0.52	0.65
SVI	High	Bottom-up	1.61	24.17	0.07	179	0.77	0.78	-0.46	0.64
SPI	Low	Top-down	1.56	23.21	0.07	264	0.43	0.03	0.13	0.44
SPI	Low	Bottom-up	1.51	23.09	0.07	297	0.43	0.13	0.17	0.44
SPI	High	Top-down	6.61	27.59	0.24	79	0.65	0.02	0.28	0.61
SPI	High	Bottom-up	2.65	23.38	0.11	149	0.62	0.16	0.37	0.65
VPI	Low	Top-down	3.10	24.65	0.13	220	0.27	0.39	0.15	0.46
VPI	Low	Bottom-up	2.18	23.33	0.09	297	0.13	0.40	0.18	0.45
VPI	High	Top-down	4.18	26.23	0.16	94	0.38	0.53	0.24	0.54
VPI	High	Bottom-up	3.23	23.78	0.14	208	0.20	0.52	0.23	0.57
SV	Low	Top-down	1.10	23.63	0.05	341	0.62	0.59	-0.32	0.16
SV	Low	Bottom-up	0.69	23.32	0.03	297	0.61	0.58	-0.29	0.14
SV	High	Top-down	2.91	26.21	0.11	173	0.93	0.96	-0.50	0.27
SV	High	Bottom-up	2.34	25.21	0.09	149	0.91	0.96	-0.46	0.27
SP	Low	Top-down	2.50	22.92	0.11	254	0.42	-0.04	0.35	0.03
SP	Low	Bottom-up	2.03	22.87	0.09	297	0.43	0.04	0.36	-0.02
SP	High	Top-down	4.10	23.66	0.17	88	0.55	-0.07	0.68	0.06
SP	High	Bottom-up	3.59	23.27	0.15	149	0.57	-0.01	0.72	-0.07
SI	Low	Top-down	-0.29	23.35	-0.01	323	0.57	0.26	-0.27	0.51
SI	Low	Bottom-up	-0.20	23.40	-0.01	297	0.57	0.30	-0.25	0.51
SI	High	Top-down	1.62	24.81	0.07	133	0.86	0.34	-0.51	0.82
SI	High	Bottom-up	-0.72	23.97	-0.03	149	0.88	0.43	-0.47	0.83
VP	Low	Top-down	3.25	23.66	0.14	258	0.13	0.35	0.35	0.04
VP	Low	Bottom-up	3.04	23.41	0.13	297	0.09	0.40	0.35	0.02
VP	High	Top-down	4.17	25.03	0.17	74	0.26	0.53	0.60	0.07
VP	High	Bottom-up	5.33	23.71	0.22	149	0.21	0.58	0.59	0.04
VI	Low	Top-down	1.04	23.97	0.04	285	0.35	0.60	-0.33	0.58
VI	Low	Bottom-up	0.80	23.59	0.03	297	0.28	0.56	-0.27	0.52
VI	High	Top-down	2.80	26.75	0.10	123	0.58	0.88	-0.49	0.81
VI	High	Bottom-up	2.43	24.99	0.10	149	0.50	0.89	-0.44	0.79
PI	Low	Top-down	1.19	22.95	0.05	273	-0.04	-0.16	0.43	0.43
PI	Low	Bottom-up	1.27	22.97	0.06	297	-0.15	-0.13	0.45	0.41
PI	High	Top-down	2.63	24.18	0.11	89	0.07	-0.29	0.85	0.61
PI	High	Bottom-up	1.64	23.33	0.07	149	-0.21	-0.28	0.83	0.56

This table shows backtest results for the average exposures-matched portfolios. The column 'Factor combinations' denotes the specific factors used to construct a portfolio with while leaving the other factors uncontrolled. The used risk factors are Size ('S'), Value ('V'), Profitability ('P'), and Investment ('I'). For example, a portfolio using factor combination 'VI' uses risk factors Value and Investment while a portfolio using the factor combination 'SVP' use the risk factors Size, Value, and Profitability. The second column denotes whether the matched exposure are low or high. The third column denotes the construction technique applied. The column 'Ann. active return' denotes a portfolio's annualized mean return in excess to the benchmark's annualized mean return. The column 'Ann. active risk' denotes the annualized volatility of a portfolio's active return. The column 'Information Ratio' contains for each portfolio the Information Ratio which is calculated by dividing a portfolio's active return by its active risk. The seventh column denotes the average number of companies a portfolio holds. The eighth, ninth, tenth, and eleventh column denote a portfolio's average factor exposure towards the risk factors Size, Value, Profitability, and Investment respectively.

Table 8: Results of the hypothesis tests

Sharpe Ratios			
Factor combination	Exposure level	Δ	p-value
SVPI	L	0.02	0.730
SVPI	H	0.03	0.758
SVP	L	0.01	0.862
SVP	H	0.05	0.688
SVI	L	0.01	0.996
SVI	H	0.01	0.877
SPI	L	0.00	0.909
SPI	H	0.00	0.853
VPI	L	0.01	0.828
VPI	H	0.09	0.493
SV	L	0.01	0.484
SV	H	0.01	0.785
SP	L	0.04	0.338
SP	H	0.01	0.868
SI	L	0.00	0.819
SI	H	0.10	0.070
VP	L	0.00	0.805
VP	H	0.15	0.050
VI	L	0.00	0.807
VI	H	0.05	0.604
PI	L	0.00	0.793
PI	H	0.04	0.643

Information Ratios			
Factor combination	Exposure level	Δ	p-value
SVPI	L	0.04	0.218
SVPI	H	0.05	0.564
SVP	L	0.03	0.484
SVP	H	0.09	0.586
SVI	L	0.02	0.454
SVI	H	0.03	0.375
SPI	L	0.00	0.975
SPI	H	0.13	0.255
VPI	L	0.04	0.505
VPI	H	0.02	0.884
SV	L	0.02	0.302
SV	H	0.02	0.735
SP	L	0.02	0.512
SP	H	0.02	0.791
SI	L	0.00	0.833
SI	H	0.10	0.046
VP	L	0.01	0.702
VP	H	0.05	0.202
VI	L	0.01	0.690
VI	H	0.00	0.981
PI	L	0.01	0.896
PI	H	0.04	0.671

The table on the left contains for all backtested average exposures-matched bottom-up and top-down portfolios the p-values for the hypothesis test with the null hypothesis that the difference between the Sharpe Ratios of the backtested top-down and bottom-up portfolios is equal to zero (H_0:Δ=0). The table on the right shows for the also shows p-values for a similar hypothesis test but for the Information Ratios. The columns 'Factor combinations' denote the specific factors used to construct a portfolio with while leaving the other factors uncontrolled. The used risk factors are Size ('S'), Value ('V'), Profitability ('P'), and Investment ('I'). For example, a portfolio using factor combination 'VI' uses risk factors Value and Investment while a portfolio using the factor combination 'SVP' use the risk factors Size, Value, and Profitability. The columns 'Δ' contain the observed difference in Sharpe Ratios and Information Ratios between the average-exposure matched top-down and bottom-up constructed portfolio. The columns 'p-value' contain the p-values for the hypothesis tests.

5.3 Enhanced top-down constructed multi-factor portfolios

The backtested results so far have shown that, when applied in a naive manner, the bottom-up construction technique seems to be superior to the bottom-up construction technique. This is in line with findings of Clarke et al. (2016), Bender and Wang (2016), Fitzgibbons et al. (2017), Leippold and Rueegg (2018), and Blitz and Vidojevic (2019). On the other hand, the results also show that when portfolios are being average exposures-matched the portfolios perform similarly in terms of Sharpe Ratio and Information Ratio despite the construction technique applied. This is in line with findings of Blitz and Vidojevic (2019). This is also in line with the findings of Leippold and Rueegg (2018) who could not find evidence that Sharpe Ratios of bottom-up constructed portfolios are statistically significant different from top-down constructed portfolios. Next, we will discuss the results of the backtested enhanced top-down portfolios.

Proponents of the bottom-up approach argue that this approach is superior since it avoids stocks having unfavorable characteristics towards factors and therefore could achieve higher factor exposures than top-down constructed portfolios. The results of our backtested naively constructed portfolios seem to be in line with this statement.

Tables 9, 10, and 11 present the results for the backtested enhanced top-down constructed portfolios (i.e. naively constructed top-down portfolios having stocks with a specified number of negative exposures removed). For the 10% portfolios, there are no backtesting results for single-factor portfolios containing stocks with only having positive exposure to four factors for the simple reason that at some moments of rebalancing the constraints are too tight to find any stocks to construct a portfolio with. The table shows that for each threshold, the average top-down constructed portfolios combining single-factor portfolios that contain only stocks having positive exposure towards two or more of the considered risk factors have just slightly less exposures to the risk factors compared the average naively constructed bottom-up portfolio. For example, the average naively constructed bottom-up 50% portfolio has factor exposures of 0.36, 0.34, 0.05, and 0.31 towards the risk factors Size, Value, Profitability, and Investment respectively. The average enhanced top-down 50% portfolio containing stocks with two or more positive factor exposures has factor exposures of 0.31, 0.31, 0.02, and 0.30 towards the risk factors Size, Value, Profitability, and Investment respectively and a Sharpe Ratio of 0.63. Another example is the average naively constructed bottom-up 10% portfolio that has factor exposures of 0.72, 0.66, 0.24, and 0.63 towards the risk factors Size, Value, Profitability, and Investment respectively. The average enhanced top-down 10% portfolio containing stocks with two or more positive factor exposures has factor exposures of 0.66, 0.48, 0.01, and 0.58 towards the risk factors Size, Value, Profitability, and Investment respectively and a Sharpe Ratio of 0.68. Only in the second example, the enhanced portfolio has clearly a higher exposure towards the Profitability factor.

More interestingly, the table also shows that in case only stocks having 3 or 4 positive factor exposures are used, top-down constructed portfolios achieve even higher factor exposures compared to their bottom-up equivalents. For example, the average naively constructed bottom-up 50% portfolio has factor exposures of 0.36, 0.34, 0.05, and 0.31 towards the risk factors Size, Value, Profitability, and Investment respectively and has a Sharpe Ratio of 0.63. The average enhanced top-down 50% portfolio containing only stocks with three or four positive factor exposures has factor exposures of 0.55, 0.59, 0.02, and 0.48 towards the risk factors Size, Value, Profitability, and Investment respectively and has a Sharpe Ratio of 0.71. Another example is the average naively constructed bottom-up 10% portfolio that has factor exposures of 0.72, 0.66, 0.24, and 0.63 towards the risk factors Size, Value, Profitability, and Investment respectively and a Sharpe Ratio of 0.71. The average enhanced top-down 10% portfolio containing stocks with three or four positive factor exposures has factor exposures of 0.88, 0.82, 0.03, and 0.76 towards the risk factors Size, Value, Profitability, and Investment respectively and has a similar Sharpe Ratio of 0.70.

The results of the backtested enhanced portfolios seem to suggest that bottom-up constructed portfolios are not necessarily more effective in capturing high factor exposures than top-down constructed portfolios. Furthermore, these results also seem to contradict the statement of Ghayur et al. (2018) that top-down constructed portfolios are limited in the factor exposures achievable compared to bottom-up constructed portfolios. These finding is in line with the finding of Blitz and Vidojevic (2019). The authors find that if one removes stocks with unfavorable characteristics towards more risk factors from top-down constructed portfolios, the more and more the portfolios become similar to bottom-up constructed portfolios.

Table 9: Backtested results of enhanced constructed top-down portfolios with threshold of 50%

Top-down equally-weighted 50% single-factor portfolios Containing stocks having positive exposure to 2 factor or more								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	9.50	16.00	0.59	353.40	0.55	0.55	-0.23	0.21
SP	9.90	15.20	0.65	378.60	0.33	0.14	0.23	0.14
SI	8.40	15.50	0.55	383.10	0.46	0.30	-0.16	0.43
VP	10.80	15.60	0.69	380.90	0.16	0.32	0.21	0.16
VI	9.30	15.80	0.59	372.10	0.29	0.47	-0.18	0.45
PI	9.80	15.20	0.64	362.10	0.07	0.06	0.27	0.38
SVP	10.10	15.50	0.65	413.80	0.35	0.34	0.07	0.17
SVI	9.10	15.70	0.58	406.60	0.43	0.44	-0.19	0.37
SPI	9.40	15.20	0.62	411.00	0.29	0.17	0.11	0.32
VPI	10.00	15.50	0.65	412.90	0.17	0.28	0.10	0.33
SVPI	9.60	15.40	0.62	415.30	0.31	0.31	0.02	0.30
Avg	**9.63**	**15.51**	**0.62**	**389.98**	**0.31**	**0.31**	**0.02**	**0.30**

Top-down equally-weighted 50% single-factor portfolios Containing stocks having positive exposure to 3 or 4 factors								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	11.00	16.20	0.68	195.20	0.66	0.70	-0.17	0.48
SP	11.50	15.50	0.74	196.20	0.53	0.48	0.22	0.39
SI	10.60	16.20	0.65	194.20	0.65	0.64	-0.17	0.58
VP	11.90	15.70	0.76	196.10	0.44	0.54	0.21	0.38
VI	10.90	16.30	0.67	194.00	0.57	0.69	-0.18	0.58
PI	11.50	15.70	0.73	191.00	0.44	0.48	0.21	0.49
SVP	11.40	15.70	0.73	197.40	0.54	0.58	0.09	0.42
SVI	10.80	16.20	0.67	197.20	0.63	0.68	-0.17	0.55
SPI	11.20	15.70	0.71	197.50	0.54	0.53	0.09	0.49
VPI	11.40	15.80	0.72	197.30	0.48	0.57	0.08	0.48
SVPI	11.20	15.80	0.71	197.50	0.55	0.59	0.02	0.48
Avg	**11.22**	**15.89**	**0.71**	**195.78**	**0.55**	**0.59**	**0.02**	**0.48**

Top-down equally-weighted 50% single-factor portfolios Containing stocks having positive exposure to 4 factors								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	15.40	18.80	0.82	30.80	0.66	0.65	0.64	0.60
SP	16.00	18.90	0.85	30.90	0.66	0.65	0.64	0.60
SI	16.00	19.00	0.84	30.80	0.67	0.65	0.65	0.62
VP	15.40	18.80	0.82	30.90	0.65	0.65	0.64	0.60
VI	15.40	18.90	0.82	30.80	0.66	0.66	0.64	0.63
PI	16.00	19.00	0.84	30.90	0.66	0.65	0.65	0.63
SVP	15.60	18.80	0.83	30.90	0.66	0.65	0.64	0.60
SVI	15.60	18.90	0.83	30.80	0.66	0.65	0.64	0.62
SPI	16.00	19.00	0.84	30.90	0.66	0.65	0.65	0.62
VPI	15.60	18.90	0.83	30.90	0.66	0.65	0.64	0.62
SVPI	15.70	18.90	0.83	30.90	0.66	0.65	0.64	0.61
Avg	**15.70**	**18.90**	**0.83**	**30.86**	**0.66**	**0.65**	**0.64**	**0.61**

This table shows the results of backtested naively top-down constructed portfolios using single-factor portfolios with a thresholds of 50%. The upper table shows the results for top-down constructed portfolios containing only stocks with two or more positive factor scores, the middle table shows the results for the top-down constructed portfolios containing only stocks with three or four positive factor scores, and the lower table shows the results for top-down constructed portfolios containing only stocks with four positive factor scores. Each table's last row contain for each column the average value in bold letters. The column 'Factor combinations' denotes the specific factors used to construct a portfolio with. The used risk factors are Size ('S'), Value ('V'), Profitability ('P'), and Investment ('I'). For example, a portfolio using factor combination 'VI' uses risk factors Value and Investment while a portfolio using the factor combination 'SVP' use the risk factors Size, Value, and Profitability. The column 'Mean excess return' denotes a portfolio's annualized mean return in excess to the risk-free rate. The column 'Excess risk' denotes the annualized volatility of a portfolio's returns in excess to the risk-free rate. The column 'Sharpe Ratio' contains for each portfolio the Sharpe Ratio which is calculated by dividing a portfolio's mean excess return by its excess risk. The fifth column denotes the average number of companies a portfolio holds. The sixth, seventh, eighth, and ninth column denote a portfolio's average factor exposure towards the risk factors Size, Value, Profitability, and Investment respectively.

Table 10: Backtested results of enhanced constructed top-down portfolios with threshold of 33.33%

Top-down equally-weighted 33.33% single-factor portfolios Containing stocks having positive exposure to 2 factor or more								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	10.30	16.80	0.62	265.70	0.74	0.71	-0.34	0.26
SP	10.30	15.10	0.68	281.00	0.44	0.11	0.33	0.17
SI	8.60	15.80	0.54	281.70	0.62	0.34	-0.25	0.57
VP	11.80	15.80	0.75	292.10	0.21	0.37	0.30	0.18
VI	10.10	16.40	0.62	281.60	0.40	0.61	-0.29	0.58
PI	10.00	15.10	0.66	271.10	0.10	0.01	0.38	0.49
SVP	10.80	15.70	0.69	346.90	0.46	0.40	0.10	0.20
SVI	9.70	16.30	0.60	327.40	0.59	0.55	-0.29	0.47
SPI	9.60	15.20	0.64	342.00	0.39	0.16	0.15	0.41
VPI	10.70	15.60	0.68	353.30	0.24	0.33	0.13	0.42
SVPI	10.20	15.60	0.65	377.90	0.42	0.36	0.02	0.38
Avg	**10.19**	**15.76**	**0.65**	**310.97**	**0.42**	**0.36**	**0.02**	**0.38**

Top-down equally-weighted 33.33% single-factor portfolios Containing stocks having positive exposure to 3 or 4 factors								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	12.00	17.30	0.70	156.50	0.83	0.87	-0.28	0.53
SP	12.10	15.50	0.78	149.40	0.65	0.49	0.33	0.43
SI	11.00	17.00	0.64	157.10	0.79	0.69	-0.24	0.71
VP	13.00	15.90	0.81	157.80	0.50	0.63	0.30	0.42
VI	11.90	17.30	0.69	158.20	0.65	0.83	-0.27	0.71
PI	11.90	15.80	0.75	143.20	0.46	0.45	0.34	0.60
SVP	12.40	16.10	0.77	178.20	0.66	0.66	0.11	0.46
SVI	11.60	17.10	0.68	175.20	0.76	0.79	-0.27	0.65
SPI	11.70	15.90	0.73	176.10	0.63	0.55	0.14	0.58
VPI	12.20	16.20	0.76	179.20	0.54	0.64	0.12	0.58
SVPI	12.00	16.30	0.74	187.80	0.65	0.66	0.03	0.57
Avg	**11.98**	**16.40**	**0.73**	**165.34**	**0.65**	**0.66**	**0.03**	**0.57**

Top-down equally-weighted 33.33% single-factor portfolios Containing stocks having positive exposure to 4 factors								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	16.60	21.30	0.78	25.70	0.81	0.80	0.65	0.64
SP	13.70	20.20	0.68	26.00	0.78	0.65	0.82	0.64
SI	16.10	21.00	0.77	25.80	0.81	0.68	0.69	0.75
VP	16.70	20.90	0.80	26.70	0.67	0.77	0.78	0.63
VI	18.90	21.90	0.86	25.40	0.69	0.81	0.65	0.75
PI	16.20	20.70	0.78	25.60	0.67	0.66	0.82	0.74
SVP	15.70	20.60	0.76	28.60	0.76	0.74	0.75	0.64
SVI	17.30	21.30	0.81	28.20	0.77	0.76	0.66	0.71
SPI	15.40	20.50	0.75	28.70	0.75	0.66	0.78	0.71
VPI	17.30	21.00	0.82	28.80	0.68	0.74	0.75	0.71
SVPI	16.40	20.80	0.79	29.70	0.74	0.73	0.73	0.69
Avg	**16.39**	**20.93**	**0.78**	**27.20**	**0.74**	**0.73**	**0.73**	**0.69**

This table shows the results of backtested naively top-down constructed portfolios using single-factor portfolios with a thresholds of 33.33%. The upper table shows the results for top-down constructed portfolios containing only stocks with two or more positive factor scores, the middle table shows the results for the top-down constructed portfolios containing only stocks with three or four positive factor scores, and the lower table shows the results for top-down constructed portfolios containing only stocks with four positive factor scores. Each table's last row contain for each column the average value in bold letters. The column 'Factor combinations' denotes the specific factors used to construct a portfolio with while leaving the other factors uncontrolled. The used risk factors are Size ('S'), Value ('V'), Profitability ('P'), and Investment ('I'). For example, a portfolio using factor combination 'VI' uses risk factors Value and Investment while a portfolio using the factor combination 'SVP' use the risk factors Size, Value, and Profitability. The column 'Mean excess return' denotes a portfolio's annualized mean return in excess to the risk-free rate. The column 'Excess risk' denotes the annualized volatility of a portfolio's returns in excess to the risk-free rate. The column 'Sharpe Ratio' contains for each portfolio the Sharpe Ratio which is calculated by dividing a portfolio's mean excess return by its excess risk. The fifth column denotes the average number of companies a portfolio holds. The sixth, seventh, eighth, and ninth column denote a portfolio's average factor exposure towards the risk factors Size, Value, Profitability, and Investment respectively.

Table 11: Backtested results of enhanced constructed top-down portfolios with threshold of 10%

Top-down equally-weighted 10% single-factor portfolios Containing stocks having positive exposure to 2 factor or more								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	13.80	21.50	0.64	93.20	1.11	1.10	-0.57	0.39
SP	12.50	16.10	0.77	93.40	0.67	0.11	0.55	0.25
SI	11.20	18.70	0.60	100.40	0.95	0.44	-0.53	0.94
VP	12.90	17.80	0.73	95.60	0.37	0.53	0.56	0.22
VI	11.80	19.70	0.60	103.50	0.64	0.85	-0.53	0.91
PI	10.40	15.70	0.66	92.10	0.21	-0.13	0.60	0.77
SVP	13.10	17.90	0.73	130.20	0.72	0.58	0.18	0.29
SVI	12.30	19.60	0.63	131.70	0.90	0.80	-0.54	0.75
SPI	11.40	16.20	0.70	134.90	0.61	0.14	0.21	0.66
VPI	11.70	17.20	0.68	139.30	0.41	0.42	0.21	0.64
SVPI	12.20	17.40	0.70	165.80	0.66	0.48	0.01	0.58
Avg	**12.12**	**17.98**	**0.68**	**116.37**	**0.66**	**0.48**	**0.01**	**0.58**

Top-down equally-weighted 10% single-factor portfolios Containing stocks having positive exposure to 3 or 4 factors								
Factor combination	Mean excess return	Excess risk	Sharpe Ratio	Avg. no. of companies	Avg. Size exposure	Avg. Value exposure	Avg. Profitability exposure	Avg. Investment exposure
SV	15.60	23.70	0.66	62.50	1.17	1.25	-0.53	0.62
SP	14.40	18.20	0.79	51.50	0.94	0.56	0.56	0.47
SI	13.40	22.50	0.60	62.40	1.08	0.79	-0.48	1.05
VP	14.80	18.80	0.79	55.90	0.68	0.85	0.55	0.47
VI	13.90	22.80	0.61	66.10	0.82	1.08	-0.50	1.04
PI	12.60	18.80	0.67	46.70	0.60	0.39	0.60	0.89
SVP	15.00	19.60	0.77	76.10	0.93	0.88	0.19	0.52
SVI	14.30	22.60	0.64	82.10	1.02	1.04	-0.50	0.90
SPI	13.50	19.10	0.71	74.60	0.87	0.58	0.23	0.80
VPI	13.80	19.40	0.71	78.90	0.70	0.77	0.22	0.80
SVPI	14.20	19.80	0.71	93.90	0.88	0.82	0.03	0.76
Avg	**14.14**	**20.48**	**0.70**	**68.25**	**0.88**	**0.82**	**0.03**	**0.76**

This table shows the results of backtested naively top-down constructed portfolios using single-factor portfolios with a thresholds of 10%. The upper table shows the results for top-down constructed portfolios containing only stocks with two or more positive factor scores, the middle table shows the results for the top-down constructed portfolios containing only stocks with three or four positive factor scores, and the lower table shows the results for top-down constructed portfolios containing only stocks with four positive factor scores. Each table's last row contain for each column the average value in bold letters. The column 'Factor combinations' denotes the specific factors used to construct a portfolio with while leaving the other factors uncontrolled. The used risk factors are Size ('S'), Value ('V'), Profitability ('P'), and Investment ('I'). For example, a portfolio using factor combination 'VI' uses risk factors Value and Investment while a portfolio using the factor combination 'SVP' use the risk factors Size, Value, and Profitability. The column 'Mean excess return' denotes a portfolio's annualized mean return in excess to the risk-free rate. The column 'Excess risk' denotes the annualized volatility of a portfolio's returns in excess to the risk-free rate. The column 'Sharpe Ratio' contains for each portfolio the Sharpe Ratio which is calculated by dividing a portfolio's mean excess return by its excess risk. The fifth column denotes the average number of companies a portfolio holds. The sixth, seventh, eighth, and ninth column denote a portfolio's average factor exposure towards the risk factors Size, Value, Profitability, and Investment respectively.

6 Conclusion

This paper backtests and compares two multi-factor portfolio construction techniques in fifteen European stock markets over the time period July 2004 till December 2022. This is done by constructing three types of portfolios. First, we backtested and analyzed naively constructed bottom-up and top-down portfolios. Secondly, we backtested and analyzed average exposures-matched bottom-up and top-down portfolios. Here, for each factor combination we construct two bottom-up portfolios using a threshold of 50% and 25%, and construct top-down portfolios matching the bottom-up portfolios' average exposures. Thirdly, we backtested enhanced constructed top-down portfolios (naively constructed top-down portfolios having stocks with a specified number of negative exposures removed) and compare these to the naively constructed bottom-up portfolios.

The backtested results seem to support the claims that, when applied in a naive manner, bottom-up constructed portfolios do outperform top-down constructed portfolios on absolute and risk-adjusted basis. According to Blitz and Vidojevic (2019) this is in line with what could be expected since the bottom-up constructed portfolios are more concentrated and therefore more effective in capturing higher factor exposures. However, it should also be noted that Leippold and Rueegg (2018) do no find statistical evidence for the outperformance of naively constructed bottom-up portfolios.

The results of the backtested average exposures-matched portfolios do not seem to support the claim that bottom-up portfolios outperform top-down portfolios. We find that none of backtested average exposures-matched top-down and bottom-up portfolios have a difference in both the Sharpe Ratio and the Information Ratio that is statistically significant different from zero at a 5% significance level. This is in line with findings of Ghayur et al. (2018) and Blitz and Vidojevic (2019) who argue that the bottom-up portfolios do not perform better than top-down portfolios when the portfolios (on average) have equal factor exposures. These results are also in line with the aforementioned findings of Leippold and Rueegg (2018). Blitz and Vidojevic (2019) argues that there does not seem to exist such a thing as a 'bottom-up premium'. That is, one could not gain higher risk-adjusted returns than expected by the factor exposures when applying a bottom-up constructed portfolio. These results seem to support this claim.

Lastly, we find that enhanced top-down constructed portfolios that only contain stocks having positive exposure towards 3 or 4 risk factors do outperform the naively constructed bottom-up portfolios. This also implies that bottom-up constructed portfolios are not necessarily more effective in capturing higher factor exposures than top-down constructed portfolios. Furthermore, this also implies that top-down constructed portfolios are not limited in the achievable factor exposures compared to bottom-up portfolios, as noted by Ghayur et al. (2018). The backtested results in this paper seem to support this claim.

The backtesting results imply that in case the investor wants to apply one of the construction techniques only in a naive manner, the bottom-up approach seem indeed to be the better choice. But as noted by Blitz and Vidojevic (2019), that is to be expected since naively constructed bottom-up portfolios are more concentrated and therefore more effective in achieving higher factor exposures. The backtesting results also seem to suggest that, in case the investor wants to construct a portfolio having a desired factor exposure, both construction techniques could be used. Top-down constructed portfolios could achieve similar or even better risk exposures compared to naively constructed bottom-up portfolios by tightening the thresholds of the single-factor portfolio and/or by removing stocks having unfavorable characteristics towards multiple factors from the single-factor portfolios. Therefore, the backtesting results seem to suggest that there is no performance difference between both portfolio construction techniques when matched at desired (average) factor exposures.

6.1 Limitations and suggestions for further research

We used the risk factors of the five-factor model proposed by Fama and French (2015) to construct multi-factor portfolios with. However, there are more risk factor found in the stock market like for example Momentum, Low-Volatility, Quality (closely related to Profitability), and Betting-against-Beta. Further research could focus on how bottom-up and top-down constructed portfolios perform compared to each other when portfolios are constructed using other risk factors. Furthermore, the stocks being used in the backtested portfolios are not filtered on for example stock price, market capitalization, or liquidity. Therefore, the backtested portfolios also contain illiquid penny stocks and illiquid micro cap stocks. Therefore, the backtested portfolios are more of a theoretical nature than of practical nature. For example, for large funds it is impossible to allocate an equal weight to micro cap companies while smaller funds may not always want to invest in illiquid stocks. Further research could investigate whether the findings in this paper also hold for portfolios containing stocks that are filtered on for example stock price, market capitalization, or liquidity.

References

Aharoni, G., Grundy, B., & Zeng, Q. (2013). Stock returns and the miller modigliani valuation formula: Revisiting the fama french analysis. *Journal of Financial Economics*, *110*(2), 347–357. https://doi.org/https://doi.org/10.1016/j.jfineco.2013.08.003

Amenc, N., Goltz, F., & Sivasubramanian, S. (2018). Multifactor index construction: A skeptical appraisal of bottom-up approaches. *The Journal of Index Investing*, *9*(1), 6–17. https://doi.org/10.3905/jii.2018.9.1.006

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, *59*(3), 817–858. Retrieved May 18, 2023, from http://www.jstor.org/stable/2938229

Andrews, D. W. K., & Monahan, J. C. (1992). An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica*, *60*(4), 953–966. Retrieved May 18, 2023, from http://www.jstor.org/stable/2951574

Ang, A. (2010). The four benchmarks of sovereign wealth funds. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1680485

Ardia, D., & Boudt, K. (2018). The peer performance ratios of hedge funds. *Journal of Banking Finance*, *87*, 351–368. https://doi.org/https://doi.org/10.1016/j.jbankfin.2017.10.014

Asness, C. (2016). The siren song of factor timing aka "smart beta timing" aka "style timing". *The Journal of Portfolio Management*, *42*, 1–6. https://doi.org/10.3905/jpm.2016.42.5.001

Asness, C., Chandra, S., Ilmanen, A., & Israel, R. (2017). Contrarian factor timing is deceptively difficult. *The Journal of Portfolio Management*, *43*(5), 72–87. https://doi.org/10.3905/jpm.2017.43.5.072

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. *The Journal of Finance*, *68*(3), 929–985. https://doi.org/https://doi.org/10.1111/jofi.12021

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics*, *9*(1), 3–18. https://doi.org/https://doi.org/10.1016/0304-405X(81)90018-0

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

Journal of Finance, *32*(3), 663–682. https://doi.org/https://doi.org/10.1111/j.1540-6261.1977.tb01979.x

Bender, J., Briand, R., Nielsen, F., & Stefek, D. (2009). Portfolio of risk premia: A new approach to diversification. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2543991

Bender, J., & Wang, T. (2016). Can the whole be more than the sum of the parts? bottom-up versus top-down multifactor portfolio construction. *The Journal of Portfolio Management*, *42*, 39–50. https://doi.org/10.3905/jpm.2016.42.5.039

Bessler, W., Taushanov, G., & Wolff, D. (2021). Factor investing and asset allocation strategies: A comparison of factor versus sector optimization. *Journal of Asset Management*, *22*. https://doi.org/10.1057/s41260-021-00225-1

Bhandari, L. C. (1988). Debt/equity ratio and expected common stock returns: Empirical evidence. *The Journal of Finance*, *43*(2), 507–528. Retrieved April 19, 2023, from http://www.jstor.org/stable/2328473

Blitz, D. (2012). Strategic allocation to premiums in the equity market. *The Journal of Beta Investment Strategies*, *2*(4), 42–49. https://doi.org/10.3905/jii.2012.2.4.042

Blitz, D. (2015). Factor investing revisited. *The Journal of Beta Investment Strategies*, *6*(2), 7–17. https://doi.org/10.3905/jii.2015.6.2.007

Blitz, D., & Vidojevic, M. (2019). The characteristics of factor investing. *The Journal of Portfolio Management*, *45*(3), 69–86. https://doi.org/10.3905/jpm.2019.45.3.069

Brière, M., & Szafarz, A. (2021). When it rains, it pours: Multifactor asset management in good and bad times. *Journal of Financial Research*, *44*(3), 641–669. https://doi.org/https://doi.org/10.1111/jfir.12257

Chow, T.-M., Li, F., & Shim, Y. (2018). Smart beta multifactor construction methodology: Mixing versus integrating. *The Journal of Index Investing*, *8*(4), 47–60. https://doi.org/10.3905/jii.2018.8.4.047

Clarke, R., de Silva, H., & Thorley, S. (2016). Fundamentals of efficient factor investing. *Financial Analysts Journal*, *72*, 1–18. https://doi.org/10.2469/faj.v72.n6.3

Cochrane, J. H. (2011). Presidential address: Discount rates. *The Journal of Finance*, *66*(4), 1047–1108. https://doi.org/https://doi.org/10.1111/j.1540-6261.2011.01671.x

Cohen, R. B., Gompers, P. A., & Vuolteenaho, T. (2002). Who underreacts to cash-flow news? evidence from trading between individuals and institutions [Limits on Arbitrage]. *Journal of Financial Economics*, *66*(2), 409–462. https://doi.org/https://doi.org/10.1016/S0304-405X(02)00229-5

Dichtl, H., Drobetz, W., Lohre, H., Rother, C., & Voßkamp, P. (2019). Optimal timing and tilting of equity factors. *Financial Analysts Journal*, *75*, 1. https://doi.org/10.1080/0015198X.2019.1645478

Dichtl, H., Drobetz, W., & Wendt, V.-S. (2020). How to build a factor portfolio: Does the allocation strategy matter? *European Financial Management*, *27*. https://doi.org/10.1111/eufm.12264

Fairfield, P., Whisenant, S., & Yohn, T. (2002). Accrued earnings and growth: Implications for future profitability and market mispricing. *The Accounting Review*, *78*. https://doi.org/10.2308/accr.2003.78.1.353

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/https://doi.org/10.1016/0304-405X(93)90023-5

Fama, E. F., & French, K. R. (2006). Profitability, investment and average returns. *Journal of Financial Economics*, *82*(3), 491–518. https://doi.org/https://doi.org/10.1016/j.jfineco.2005.09.009

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/https://doi.org/10.1016/j.jfineco.2014.10.010

Fitzgibbons, S., Friedman, J., Pomorski, L., & Serban, L. (2017). Long-only style investing: Don't just mix, integrate. *The Journal of Investing*, *26*(4), 153–164. https://doi.org/10.3905/joi.2017.26.4.153

Ghayur, K., Heaney, R., & Platt, S. (2018). Constructing long-only multifactor strategies: Portfolio blending vs. signal blending. *Financial Analysts Journal*, *74*, 1–17. https://doi.org/10.2469/faj.v74.n3.5

Harvey, C. R., Liu, Y., & Zhu, H. (2016). ... and the cross-section of expected returns. *The Review of Financial Studies*, *29*(1), 5–68. Retrieved February 8, 2023, from http://www.jstor.org/stable/43866011

Hasan, M., Alam, M., Amin, M., & Rahaman, M. (2015). The size and value effect to explain cross-section of expected stock returns in dhaka stock exchange. *International Journal of Economics and Finance*, *7*. https://doi.org/10.5539/ijef.v7n1p14

Haugen, R., & Baker, N. (1996). Commonality in the determinants of expected stock returns. *Journal of Financial Economics*, *41*, 401–439. https://doi.org/10.1016/0304-405X(95)00868-F

Haugen, R. A., & Heins, A. J. (1972). On the evidence supporting the existence of risk premiums in the capital market. *ERN: Econometric Modeling in Financial Economics (Topic)*.

Hjalmarsson, E. (2009). Diversification across characteristics. (986). https://ideas.repec.org/p/fip/fedgif/986.html

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. Retrieved May 19, 2023, from http://www.jstor.org/stable/4615733

Ilmanen, A., & Kizer, J. (2012). The death of diversification has been greatly exaggerated. *The Journal of Portfolio Management*, *38*(3), 15–27. https://doi.org/10.3905/jpm.2012.38.3.015

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, *48*(1), 65–91. Retrieved February 8, 2023, from http://www.jstor.org/stable/2328882

Jobson, J. D., & Korkie, B. M. (1981). Performance hypothesis testing with the sharpe and treynor measures. *The Journal of Finance*, *36*(4), 889–908. https://doi.org/https://doi.org/10.1111/j.1540-6261.1981.tb04891.x

Ledoit, O., & Wolf, M. (2007). Robust performance hypothesis tests with the sharpe ratio. *Journal of Empirical Finance*, *15*, 850–859. https://doi.org/10.1016/j.jempfin.2008.03.002

Leippold, M., & Rueegg, R. (2018). The mixed vs the integrated approach to style investing: Much ado about nothing? *European Financial Management*, *24*(5), 829–855. https://doi.org/https://doi.org/10.1111/eufm.12139

Lester, A. (2019). On the theory and practice of multifactor portfolios. *The Journal of Portfolio Management*, *45*(3), 87–100. https://doi.org/10.3905/jpm.2019.45.3.087

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, *47*(1), 13–37. Retrieved February 6, 2023, from http://www.jstor.org/stable/1924119

Markowitz, H. (1952). Portfolio selection*. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/https://doi.org/10.1111/j.1540-6261.1952.tb01525.x

Memmel, C. (2003). Performance hypothesis testing with the sharpe ratio. *Finance Letters*, *1*.

Miller, M., & Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares. *The Journal of Business*, *34*, 411–411. https://doi.org/10.1086/294442

Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica*, *34*(4), 768–783. Retrieved April 16, 2023, from http://www.jstor.org/stable/1910098

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, *108*(1), 1–28. https://doi.org/https://doi.org/10.1016/j.jfineco.2013.01.003

Reinganum, M. R. (1981). Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values. *Journal of Financial Economics*, *9*(1), 19–46. https://doi.org/https://doi.org/10.1016/0304-405X(81)90019-2

Richardson, S., & Sloan, R. (2003). External financing and future stock returns. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.383240

Rosenberg, B., Reid, K., & Lanstein, R. (1985). Persuasive evidence of market inefficiency. *The Journal of Portfolio Management*, *11*(3), 9–16. https://doi.org/10.3905/jpm.1985.409007

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk*. *The Journal of Finance*, *19*(3), 425–442. https://doi.org/https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Sharpe, W. F. (1994). The sharpe ratio.

Stattman, D. (1980). Book values and stock returns. *The Chicago MBA: A Journal of Selected Papers*, *4*, 25–45.

Titman, S., Xie, F., & Wei, K. (2004). Capital investments and stock returns. *Journal of Financial and Quantitative Analysis*, *39*, 677–700. https://doi.org/10.2139/ssrn.268538