Seminar Paper
at the TUM School of Management
of the Technical University of Munich

Exploring Factor-based Strategies for Portfolio Construction in the UK Stock Market

Advanced Seminar Finance & Accounting (MGT001336):
Empirical Asset Pricing

Reviewer: Prof. Dr. Christoph Kaserer
Chair of Financial Management and Capital Markets
TUM School of Management
Technology University of Munich

Advisor: Minghui Chen

Composed by:

Yuru Chen
TUM School of Management
D-80333 Munich
Matriculation No. 03750984

Tong Deng
TUM School of Management
D-80333 Munich
Matriculation No. 03743148

Yasuaki Kojima
TUM School of Management
D-80333 Munich
Matriculation No. 03769934

Yu Wang
TUM School of Management
D-80333 Munich
Matriculation No. 03763623

Ziyang Zhou
TUM School of Management
D-80333 Munich
Matriculation No. 03742353

Submitted on: July 24, 2023

Contents

List of Abbreviations

AG Asset Growth

Alpha_3 Historical Alpha (3 Years)

B/M Book to Market Ratio

B/Mm Book to Market Ratio Using the Most Recent MV

Beta_3 Beta (3 Years)

C/P Cash-flow to price

C/Pm Cash-Flow to Price Using the Most Recent MV)

CEI Composite Equity Issuance

DY Dividend Yield

E/P Earnings to Price

E/Pm Earnings to price Using the Most Recent MV

EBITDA/EV Enterprise Multiple

EPS Historical Earnings per Share Growth Rate

GP/A Gross Profit to Asset

I/A Investment to Assets

Log_MV Log of Market Value

LSE London Stock Exchange

Mom_12_2 Momentum_12_2

MV Market Value

NOA Net Operating Asset

NSI Net Stock Issues

NYSE New York Stock Exchange

OA Operating Accrual

OP/BE Operating Profit-to-Book Equity

ROA Return on Asset

ROE Return on Equity

RSI Relevant Strength Index

SD Monthly Standard Deviation

Sigma_3 Sigma (3 Years)

SPS Historical Sales per Share Growth Rate

VAR Vector Autoregression

List of Tables

1 Introduction

1.1 The Various Schemes of Investing

With the outbreak of the Corona epidemic, the easily downloadable stock trading platform Robinhood Markets emerged into public awareness. Greedy first-time investors, in particular, flocked into the exciting world of investing after Robinhood advertised no trading fees. Clients, on the other hand, overlooked the reality that they also suffered a lot in wealth,[1] which makes it more necessary for them to discover the underlying investing philosophy.

Different investing styles are pursued by asset managers and investors, such as active investing, passive investing, value investing, growth investing, etc. Interestingly, what lies between active investing and passive investing is the well known approach of factor investing.[2]

Factor investing relies on the existence of factors that have generated a premium over extended periods of time, reflected exposure to systematic risk, and been validated by academic research. Such strategies of factor investing focus on securities with particular attributes involving quality, size, momentum, etc. Factors are rule-based and enduring features that assist investors in understanding variances in expected returns and earning active returns.[3]

1.2 Our Project and Motivation

In reality, capital markets are currently neither "fully efficient" nor "inefficient". However, with the development of technology, markets have grown more and more efficient in recent decades, as researchers have identified hundreds of factors that could explain stock returns.[4] Hence, It is critical to determine the key variables inside the factor zoo in order to implement the most lucrative and effective investment forecasting strategies.

In the following sections, we will discuss our investment strategy under the context of an inefficient market, no transaction costs, no short positions and only big stocks. The objective of this paper is to present a comprehensive analysis of our design process, investment strategy, final results, and further potential prospects.

[1] See Bisnoff (2022).
[2] See Bender et al. (2013), p. 15.
[3] See Bender et al. (2013), p. 2.
[4] See McLean and Pontiff (2016), p. 5.

2 Literature review

2.1 The Overview of UK Capital Market and Its Anomalies

The New York Stock Exchange (NYSE) and NASDAQ in the United States, the London Stock Exchange (LSE) in the United Kingdom, the Tokyo Stock Exchange in Japan, and the Shanghai Stock Exchange in China are among the major worldwide stock exchanges. Macroeconomic circumstances, geopolitical conflicts, business earnings, and investor sentiment all have an impact on the global stock market's performance, contributing to the oscillation of volatile stock markets.

The UK stock market is one of the largest and most prominent stock markets in the world, attracting both domestic and international investors and exerting significant influence on global financial markets. The LSE is the primary stock exchange in the United Kingdom. The LSE's listed firms are divided into indices such as the FTSE 100, which contains the top 100 companies by market capitalization, and the FTSE 250, which includes the next 250 wealthiest companies.

The UK stock market lures companies from all over the world, making it a hub for international listings. Numerous corporations from Europe, Asia, etc want to list their shares on the LSE in order to have access to global investors and raise financing. Finance, healthcare, consumer products, technology, energy, and other industries are all covered on the UK stock exchange. This diversity allows investors to participate in a variety of businesses and sectors based on their tastes and investing techniques.[5]

In terms of the presence and features of asset growth anomalies in the UK stock market, Slotte pointed out a negative relationship between total asset growth and predicted stock returns. Nonetheless, the observed influence was not as long-lasting and peculiar as in the US stock market. The UK portfolio revealed that the anomaly profits were strongest among large and minor corporations, whereas medium-sized businesses were not economically feasible anomalies. Past performances influenced anomaly earnings. According to the regression analysis on the individual stock level, total asset growth was an influential component of cross-sectional stock returns.[6]

2.2 Factor Construction

2.2.1 The History of FF5FM and Momentum

To begin with, we evaluated the practicality of various factor predictors in the grand equity asset pricing world. Back in 1960, the Capital Asset Pricing Model (CAPM) emerged, revolutionizing the capital market theory. It provided the most logical and

[5] See Meek and Gray (1989), p. 316.
[6] See Slotte (2021), p. 61.

essential explanation for the relationship between asset returns and market performance developed by Jack Treynor, William Sharpe, John Lintner, and Jan Mossin.[7]

However, scientists and researchers later on realized the limitations of CAPM such as the credibility of assumptions, etc, even though it has been widely used over the years. In 1992, Fama and French developed a three-factor model, in which they determined the explanatory capacity of size and book-to-market ratio for average returns on stocks cross-sectionally. The Fama and French three-factor model, one of the most popular theories attempting to resolve the anomalies, can be represented by the following equation:[8]

$$R_{i,t} = \alpha_i + \beta_{i,1}\left(R_M - R_{f,t}\right) + \beta_{i,2} SMB_t + \beta_{i,3} HML_t + \varepsilon_{i,t} \tag{1}$$

Here the left hand of the equation $R_{i,t}$ represents the return of the portfolio, as for the right hand of the equation, α_i is the intercepted item, $R_M - R_{f,t}$ is the market excess return, and the parameters $\beta_{i,1}$, $\beta_{i,2}$, $\beta_{i,3}$ measure the risk exposure to each of the three factors respectively, and finally residual error $\varepsilon_{i,t}$ comes at the end. The factor SMB means the difference between small and large market capitalization equities, whilst the factor HML denotes the difference between high and low book-to-market capitalization companies. The application of the three-factor model to evaluate asset performance has greatly evolved into industry standards since its inception. This three-factor model surpasses the CAPM model in describing equities returns in the US market. The majority of abnormalities can be attributed to these three components with the exception only of short-term momentum.[9]

With the rapid development of technology, Fama and French in 2015 revised the three-factor model by incorporating two additional factors, leading to the invention of the most famous Fama and French five-factor model, shown in the following equation.[10]

$$R_{it} - R_{Ft} = a_i + b_i(R_{Mt} - F_{Ft}) + s_i SMB_t + h_i HML + r_i RMW + c_i CMA + e_{it} \tag{2}$$

As we can see from the formula, the factors RMW and CMA are additional profitability and investment factors. RMW is the difference between the return spreads of the most profitable and the least profitable enterprises, while CMA is the return spread between conservatively and aggressively investing organizations.[11]

Other than Fama and French, Carhart proposed a generation of Fama and French 3-factor model, the Carhart 4-factor model. Carhart added a momentum element to stock asset price. Momentum is the velocity at which a stock price differs. To be specific, the monthly momentum factor is determined by deducting the equal-weighted

[7] See Sullivan (2007), p. 207.
[8] See Fama and French (1993), p. 35.
[9] See Fama and French (1993), p. 21.
[10] See Fama and French (2015), p. 1.
[11] See Fama and French (2015), p. 3.

average of the lowest-performing enterprises from the equal-weighted average of the highest-performing enterprises, both lag one month.[12]

2.2.2 Factors from the MSCI Structure

MSCI Inc. has developed a family of factor indexes to capture the performance of equity risk premia factors, creating a powerful means for investors to have access to variables in a practical and transparent manner. Factor allocations can be implemented passively using factor indexes, potentially saving institutional investors' budgets. Furthermore, factor indexes increase transparency in factor allocations, benefiting risk management.[13]

Size, momentum, yield, quality, liquidity, volatility, growth, and value constitute eight categories of the MSCI index. Size factor refers to the excess returns of smaller enterprises over their larger counterparts, usually captured by market capitalization. Momentum captures the stocks with higher past performance, represented by relative returns of 3-month, 6-month, 12-month, and historical alpha, etc. Dividend Yield tries to seize companies with higher-than-average dividend yields. In terms of quality factor, it apprehends the premium of stocks that are distinguished by low debt or other quality metrics, suggesting stable earnings growth. ROE, financial leverage, the strength of management, accruals, cash flows, etc can all be used to measure quality factors. Liquidity records variations in stock profits on account of the number of relative trading volumes. Annualized traded value ratio, monthly share turnover, and annual share turnover can capture the liquidity factors. Volatility records beta, lower than average volatility, or idiosyncratic risk of stocks, showed by 1 to 3-year standard deviation, downside standard deviation, beta, etc. Similarly to value factor, growth factor evaluates company growth visions utilizing sales growth and earnings growth. Typical variables are sales per share growth rate predicted long-term growth, and earnings per share growth rate. Finally, the value factor estimates excess returns to stocks that have low prices relative to their fundamental value. Book to price, earnings to price, book value, net profit, cash flow, etc are favorable indicators.[14]

MSCI has been very popular regarding stock index approaches. There are various papers that adopt MSCI index methods. For example, Antonelli et al. calculated benchmark effects from the Chinese stock market using the MSCI Emerging Markets index and concluded index inclusion has drastically boosted capital flows to China via mutual funds, suggesting the rising significance of benchmark-driven investment altering the asset management business.[15] The MSCI index is also useful when it comes to hedge with foreign currency in stock index futures, suggested by a paper written by Wang et al. (2003). In the context of a global capital interest, for-

[12] See Carhart (1997), p. 80.
[13] See Bender et al. (2013), p. 2.
[14] See Bender et al. (2013), p. 5.
[15] See Antonelli et al. (2022), p. 1.

eign currency-oriented derivatives could thrive, clarifying the strong liquidity in the MSCI Taiwan index futures market to some extent.[16] Not to mention MSCI has the sufficient capacity to summarize each country's capital market correlations. Jung et al. (2006) used MSCI to analyze the Korean stock market and discovered that the Korean stock market failed to identify the clustering of industrial pairs or business categories.[17]

2.2.3 Factors from Hanauer and Lauterbach, 2019

In Hanauer and Lauterbach 2019's research regarding the cross-sectional emerging market stock returns, they concentrated on essential corporate characteristics proposed in the empirical asset pricing literature and generated fruitful forecasting results in the end. For instance, the earnings-to-price ratio (E/Pm) using the most recent market value, book-to-market ratio (B/Mm) using the most recent market value, cash flow-to-price using the most recent MV (C/Pm), cash flow-to-price (C/P) are several examples for their value factor construction. Their quality factor includes gross profit to asset (GP/A), operating profit-to-book equity (OP/BE), Net operating asset (NOA), etc. In terms of investment factors, they considered asset growth (AG), net stock issues (NSI), composite equity issuance (CEI), investment to assets (I/A), etc. One thing to pinpoint is that they omitted negative values of the respective type for value and profitability factors containing profit measures such as cash flow, gross profits, or earnings in the numerator, as negative values would be problematic to comprehend.[18]

2.3 Factor Selection

There are various methods to determine factor selection from factor zoos. We started this subsection by looking at the univariate sort first.

2.3.1 Univariate Sort

Univariate sort serves as a benchmark to assess the univariate impacts of each risk variable on predicted bond returns according to Gebhardt et al. in their research about the cross-sectional expected corporate bond returns.[19] They deduced that univariate sorting analysis implied that default betas and term betas are better predictors of average bond returns rather than solely depending on their duration and ratings.[20] Likewise, in Nartea et al.'s research regarding excessive profits in emerging markets, they found out the features of the maximum daily returns by univariate sorting approach, such as high maximum daily returns stocks typically showed the tendency to have a low book-to-market ratio.[21]

---

[16] See Wang and Low (2003), p. 1.
[17] See Jung et al. (2006), p. 1.
[18] See Hanauer and Lauterbach (2019), p. 268.
[19] See Gebhardt et al. (2005), p. 12.
[20] See Gebhardt et al. (2005), p. 14.
[21] See Nartea et al. (2017), p. 8.

2.3.2 Double Sort

Back in 1993, Fama and French used time-series regression to see if their model could capture common fluctuations in stock returns related to size and book-to-market equity in the US market. They used more than 20 portfolios double-sorted on size and book-to-market ratio as the main test instruments,[22] inspiring the researchers to adopt the same method later on. Double sort gradually becomes a popular procedure when it comes to factor selection. Blackburn et al. (2017) concluded that global returns tend to be defined by both momentum and long-term reversals concurrently utilizing double sorts.[23] Similarly, Artmann et al. (2012) validated the German stock market had an immense momentum effect contrasting the non-existence of a size or book-to-market effect by implementing double sorts.[24]

2.3.3 Fama-Macbeth Regression

Fama and MacBeth explored the relationship between risk and average return for NYSE common stocks in their work in 1973. The paper proposed and tested a hypothesis that the pricing of common stocks reflects the attempts of risk-averse investors to hold portfolios that are "efficient" in terms of expected value and dispersion of return. They presented three implications: a linear relationship between the expected return of a security and its risk in any efficient portfolio, beta as a complete measure of security risk, and an expected return equation that includes a risk premium of beta times the difference between expected return and riskless return[25].

Fama and MacBeth tested the relationship between risk and average return for NYSE common stocks using the two-parameter portfolio model and models of market equilibrium, as well as a regression equation that is the time series average of the cross-sectional regression equation. The regression equation includes the explanatory variables which are the average of the beta values for securities in the portfolio, the average of the squared values of these betas, and the average of the residuals for securities in the portfolio. The regression is estimated using least-squares estimates of the stochastic coefficients. The results from the regression are then used to test the various implications of the two-parameter model[26].

The paper also mentioned the tests conducted by Friend and Blume (1970), as well as Black et al. (1972). These tests support the hypothesis that the expected return on stocks is systematically greater than the risk-free rate, indicating that the prices of high-beta securities are relatively low relative to low-beta securities[27].

Fama and MacBeth's approach proved robust in validating most of the empirical

[22] See Fama and French (1993), p. 11.
[23] See Blackburn and Cakici (2017), p. 20.
[24] See Antonelli et al. (2022), p. 41.
[25] See Fama and MacBeth (1973), p. 607.
[26] See Fama and MacBeth (1973), pp. 614–616.
[27] See Friend and Blume (1970); Jensen et al. (2006), p. 42.

findings of their paper. Given the scarcity of previous empirical work on testing market efficiency within the context of the two-parameter model[28], their methodology has become a standard in the financial literature, lauded for its simplicity and clarity.[29]

2.4 The Cross-Sectional Forecasting of Stock Returns

Anticipating stock returns is an intriguing endeavor over long-standing decades. It is not surprising that financial professionals evaluate an assortment of variables to forecast returns aiming to improve investing performances. On the other hand, academics are also particularly interested in the projection of equity returns as the essential implications for market efficiency assessments.[30]

2.4.1 Lewellen's Regression

Lewellen investigated the cross-sectional features of return projections derived from Fama-MacBeth regression. To be more specific, he used companies' present traits and history slopes from Fama-MacBeth regressions to replicate expected returns in the real world. These projections simulated how investors could integrate various firm factors in real-time to produce composite estimations of stocks' expected returns. Forecasts varied significantly between stocks empirically and had considerable predictive power for actual performance. The estimates of projected returns were based on firms' beginning-of-month attributes and either the prior 10-year rolling average or the cumulative average of past Fama-MacBeth slopes.[31]

His findings suggested that Fama-MacBeth-based return projections indeed match genuine equity expected returns, particularly over shorter durations. And the expected return estimations are highly consistent with predictive accuracy of at least a year. In an alternative perspective, the tests imply that Fama-MacBeth regressions are an adequate procedure to aggregate a variety of firm attributes into combined evaluations of stocks' predicted returns in real-time. The cross-sectional slopes appear to be reasonably steady and well approximated.[32]

2.4.2 Vector Autoregression

Vector Autoregressive (VAR) model also has well-liked applications when it comes to stock forecasting. It is a logical extension of the univariate autoregressive model to a multivariate time series, proven to be extremely beneficial for understanding and forecasting the dynamic behavior of economic and financial time series. Predictions from VAR models can be generated conditional on the likely future courses of selected variables in the framework, making them highly flexible.[33]

---

[28] See Fama and MacBeth (1973), p. 614.
[29] See Pasquariello (2000), p. 1.
[30] See Rapach and Zhou (2023), p. 1.
[31] See Lewellen (2015), p. 14.
[32] See Lewellen (2015), p. 38.
[33] See Zivot and Wang (2006), p. 385.

Basci et al. employed VAR to investigate the connection between the Turkish stock market and an ensemble of four macroeconomic variables such as gold, import, etc. They concluded all these variables have seasonal movements.[34] Similarly, Yang examined market segmentation and information asymmetry in Chinese stock markets by executing a VAR analysis. He discovered various empirical regularities and compared them with earlier research, for example, A-share markets in both Shanghai and Shenzhen are segmented with B-share markets in the long run.[35]

2.5 The Recap of Asset Allocation Techniques

Asset allocation in portfolio management is critical since it not only balances risk but also boosts earnings. Among available allocation approaches, the tangent portfolio is a portfolio of risky assets on the efficient frontier at the point where the capital market line intersects the efficiency frontier. These portfolios offer a substantial gain in mean return when compared to risk-free investing possibilities.[36] One of its key advantages is the tangency portfolio fosters diversification and risk reduction. It enables investors to reduce the overall volatility of their financial holdings by selecting a variety of asset classes with dissimilar levels of risks and returns, minimizing the influence of market volatility on their portfolios. Maximizing the Sharpe ratio is intended to optimize risk-adjusted returns or to provide the maximum potential return for a given degree of risk, especially significant for investors who have specified risk tolerance and investment objectives.[37]

To reduce portfolio turnover, Hanauer and Lauterbach built their portfolios at the end of each quarter with the optimization of the maximum Sharpe Ratio to their emerging market factor model and noticed that they outperformed other strategies.[38]Theron and Vuuren also pointed out something similar in their research about the maximum diversification investment strategy. The tangent portfolio was the best-performing portfolio combination as it was the only risk-based portfolio that continuously produced positive returns. The portfolio not only outperformed in terms of profits, but also retained low daily volatility and a decent Sharpe Ratio.[39]

3 Data and Methodology

3.1 Data Preprocess and Factor Construction

We developed our investment strategy using two types of data: internal and external. The internal data comes from the Chair of Financial Management and Capital Markets at TUM. It includes monthly, yearly, and static information about UK stocks.

[34] See Başcı and Karaca (2013), p. 167.
[35] See Yang (2003), p. 18.
[36] See Ogryczak et al. (2015), p. 3.
[37] See Choueifaty and Coignard (2008), p. 51.
[38] See Hanauer and Lauterbach (2019), p. 277.
[39] See Theron and van Vuuren (2018), p. 14.

For external data, we got the historical return of the US 6-Month Treasury Bill, which we used as a risk-free rate, from Federal Reserve Economic Data[40]. To work with the data easily, we imported everything into RStudio and merged it into a single sheet. This way, we had all the information we need to make informed investment decisions.

In our research to understand what affects returns in the UK stock market, we explored the framework from MSCI Inc[41]. This framework includes 8 categories of factors that could influence returns, such as company size, value, yield, quality, volatility, liquidity, growth, and momentum. Simultaneously, we delved into additional factors elucidated in the research by Hanauer and Lauterbach[42]. In the end, we identified and constructed 28 factors in total. They are B/m, B/Mm[43], E/P, E/Pm[44], C/P, C/Pm[45], EBITDA/EV, Log_MV, MV, RSI, Alpha_3, Mom_12_2[46], DY, ROE, ROA, GP/A, OP/BE[47], NOA[48], OA[49], Beta_3, Sigma_3, SD, EPS, SPS, AG[50], NSI[51], CEI[52], I/A. Among them, 8 factors come from the MSCI framework, 14 factors are from Hanauer and Lauterbach's paper, and 6 factors are from both of these resources. Using these factors, we aimed to build strong investment strategies and make well-informed decisions in the ever-changing UK stock market. More details of those factors can be found in the appendix.

3.2 Factor Selection

3.2.1 Univariate Sort

As previously stated, our analysis involved a total of 28 factors. To evaluate the individual impact of each factor, we employed a method called univariate sort. And we take market value as an illustrative example to explain this process:

1. Initially, we divided all stocks into 10 deciles based on their market value.

2. Next, we created a portfolio by taking a long position in the stocks belonging to the 10th decile (highest market value) and a short position in the stocks from the 1st decile (lowest market value). This enables us to calculate the return of this portfolio.

3. We then repeated the above steps for each time point in our dataset to generate the returns for various periods.

[40] Board of Governors of the Federal Reserve System (US) (2023).
[41] See MSCI Inc. (2023).
[42] See Hanauer and Lauterbach (2019).
[43] See Asness and Frazzini (2013).
[44] See Asness and Frazzini (2013).
[45] See Asness and Frazzini (2013).
[46] See Fama and French (2008); Jegadeesh and Titman (1993).
[47] See Fama and French (2015).
[48] See Hirshleifer et al. (2004).
[49] See Sloan (1996).
[50] See Cooper et al. (2008).
[51] See Pontiff and Woodgate (2008).
[52] See Daniel and Titman (2006).

4. To determine the statistical significance of these returns, we conducted a t-test, which tests whether the portfolio's return significantly deviates from zero.

Importantly, we performed two versions of univariate sort: market-value weighted and equally weighted versions. In the market-value weighted version, we applied the market-value weighted mean to compute the portfolio return in step 2. Conversely, in the equally weighted version, we used equally weight mean within the portfolio. Subsequently, we only selected factors that demonstrate statistical significance in both versions of univariate sort. After rigorous analysis, we identified and retained 13 factors from six distinct categories, showcasing their significant impact on stock returns.

MSCI Category	Abbreviation	Description	Coefficient, (P-Value)	
			Equal-Weighted	Value-Weighted
Value	B/M	Book to market ratio	0.06 (0.7)	0.22 (0.32)
	B/Mm	Book to market ratio using the most recent MV	-2.09 (<2.2e-16)	-1.47 (<2.50e-07)
	E/P	Earnings to price	0.19 (0.32)	0.16 (0.48)
	E/Pm	Earnings to price using the most recent MV	-2.31 (<2.2e-16)	-1.77 (<1.95e-11)
	C/P	Cash flow-to-price	0.32 (0.30)	-0.06 (0.85)
	C/Pm	Cash flow-to-price using the most recent MV	-1.66 (<1.06e-05)	-1.23 (0,00)
	EBITDA/EV	Enterprise Multiple	0.46 (0.02)	0.30 (0.26)
Size	Log_MV	Log of market value	3.26 (<2.2e-16)	2.73 (<2.2e-16)
	MV	Market value	3.25 (<2.2e-16)	2.73 (<2.2e-16)
Momentum	RSI	Relevant strength index	7.93 (<2.2e-16)	7.34 (<2.2e-16)
	Alpha_3	Historical alpha (3 year)	4.42 (<2.2e-16)	4.11 (<2.2e-16)
	Mom_12_2	Momentum_12_2	0.77 (0.02)	0.74 (0.03)
Yield	DY	Diviend yield	-0.34 (0.21)	-0.37 (0.21)
Quality	ROE	Return on equity	-0.39 (0.04)	-0.71 (0.00)

Continued on next page

Table 1: Univariate sort

MSCI Category	Abbreviation	Description	Coefficient, (P-Value)	
			Equal-Weighted	Value-Weighted
	ROA	Return on asset	-0.06 (0.76)	-0.25 (0.34)
	GP/A	Gross profit to asset	0.47 (0.02)	0.42 (0.09)
	OP/BE	Operating profit-to-book equity	-0.18 (0.43)	-0.60 (0.04)
	NOA	Net operating asset	-1.34 (<1.90e-05)	-1.24 (<6.11e-05)
	OA	Operating accrual	0.46 (0.05)	0.35 (0.16)
Volatility	Beta_3	Beta (3 year)	0.99 (0.03)	0.78 (0.08)
	Sigma_3	Sigma (3 year)	25.71 (<2.2e-16)	24.26 (<2.2e-16)
	SD	Monthly standard deviation	2.45 (<4.81e-08)	2.60 (<1.43e-07)
Growth	EPS	Historical earnings per share growth rate	0.05 (0.77)	-0.17 (0.41)
	SPS	Historical sales per share growth rate	-0.06 (0.73)	-0.33 (0.20)
Investment	AG	Asset growth	-0.00 (0.96)	-0.11 (0.60)
	NSI	Net stock issues	0.31 (0.10	0.20 (0.31)
	CEI	Composite equity issuance	-0.89 (<0.00)	-0.72 (<0.00)
	I/A	Investment to assets	-0.09 (0.60)	-0.21 (0.43)

Table 1: Univariate sort (Continued)

3.2.2 Double Sort

In our study, we focused on applying a double sorting approach to the Fama-French 5 factors and momentum, which encompass SMB (Small Minus Big), HML (High Minus Low), RMW (Robust Minus Weak), CMA (Conservative Minus Aggressive), and WML (Winners Minus Losers). This method entails a 2x3 double sorting process, and we use HML as an example to explain the steps involved:

1. Initially, we divided all stocks into two categories: ”big” stocks and ”small” stocks based on their market capitalization. ”Big” stocks encompass the largest stocks, whose combined market share constitutes 90% of the total market share at that specific time. Conversely, ”small” stocks are the remaining stocks after excluding the ”big” ones.

2. Within both the ”big” and ”small” stock sets, we further categorized stocks into

three groups based on their B/M ratio: ”high,” ” medium,” and ”low” stocks. ”High” stocks represent the top 30% with the highest B/M ratios, ”medium” stocks comprise the middle 40% in terms of B/M ratios, and ”low” stocks account for the bottom 30% in B/M ratios. This dual sorting process results in six distinct subsets: ”small-high,” ”small-medium,” ”small-low,” ”big-high,” ”big-medium,” and ”big-low” sets. We then computed the market-value weighted returns for each of the six subsets (S/H, S/M, S/L, B/H, B/M, B/L).

3. We calculated the value of SMB and HML using the following formular:

$$HML = \frac{S/H + S/H}{2} - \frac{S/L + S/L}{2} \tag{3}$$

$$SMB_{HML} = \frac{S/L + S/M + S/H}{3} - \frac{B/L + B/M + B/H}{3} \tag{4}$$

4. We repeated step 1 to step 3 for each time and got a time series data of SMB and HML. Then, we conducted t-test to check if HML is significantly different from 0.

5. We repeated step 1 to step 4 to calculate other factors, namely RMW, CMA, and WML, as well as SMB_{RMW}, SMB_{CMA}, and SMB_{WML}. Similar t-tests are conducted to assess the significance of these factors.

6. Finally, we computed SMB using a designated formula and subjected it to a t-test.

$$SMB = \frac{SMB_{HML} + SMB_{RMW} + SMB_{CMA} + SMB_{WML}}{4} \tag{5}$$

The result is shown in the following table:

	Mean	T Value	P Value
SMB	0.85	7.85	<2.60e-14
HML	0.19	1.39	0.16
RMW	-0.20	-1.80	0.07
CMA	0.07	0.73	0.47
WML	0.56	2.86	<4.44e-3

Table 2: Result of double sort

Based on the t-test results, we observed that both the size premium (SMB) and the momentum premium (WML) exhibit significant positive values. These findings indicate that small cap stocks outperform big cap stocks, and that past winners outperform past losers.

3.2.3 Fama-Macbeth Regression

Hanauer and Lauterbach proposed a procedure to identify explanatory variables using Fama and Macbeth regression[53]. In our work, we adopted a similar approach

[53] See Hanauer and Lauterbach (2019), pp. 272–275.

with a modified set of variables. We began by employing the univariate sorting process and categorized the variables into six groups based on the Fama-French 5-factor model plus momentum. These groups include risk, value, profitability, investment, sizer, and momentum variables. Initially, we tested Fama and French's original selection of variables and then compared our alternatives within each category.

Table 3 shows the results of the cross-sectional regressions. In Panel A, we examined the established variables from Fama[54]. The average regression slopes and p-values for the t-statistics were calculated. Interestingly, the results showed that the average regression slopes for all variables were not significant, which aligns with the findings of the univariate sort. Moving on to Panel B, we introduced Mom_12_2 to the model. Surprisingly, this addition did not significantly alter the results for the other variables, but it did reveal a high regression coefficient for momentum itself.

Panel C was dedicated to comparing the risk variables. After including sigma 3 and SD, we found that beta 3 exhibited a significant slope. We further expanded the regression by adding B/Mm. Both B/Mm and B/M (calculated with the previous annually published market value) demonstrated high average regression coefficients and low p-values. However, to address multicollinearity concerns, as suggested by Hanauer and Lauterbach, we decided to retain only one of them and dropped B/M. In the subsequent step, E/Pm and C/Pm did not exhibit high regression slopes, leading us to exclude them from the model. Panel E delved into the comparison of profitability values. Interestingly, Beta_3 lost its significance once stronger profitability variables than OP/BE were included in the regression. In Panel F, we reaffirmed the findings of the univariate sort in Table 1, showing that CEI is a more potent investment variable than AG. During the univariate sort, MV appeared to have a similar return predictive power to Log_MV. However, after controlling for other variables, Log_MV displayed a much stronger predictive power than MV in the cross-sectional regression. Finally, we introduced two additional momentum variables, Alpha_3 and RSI, to the regression. The results demonstrated that Alpha_3 and RSI had a larger explanatory power than Mom_12_2.

Consequently, we arrived at our preliminary final model, presented in Panel F. The average regression slopes for all variables were significantly different from zero, except for NOA, which was too close to zero. As a result, we decided to omit NOA to achieve a more concise factor construction. To address concerns that the results might be driven by small stocks, we conducted a robustness check using only big stocks, presented in Panel H. Remarkably, the main results for big stocks were identical, leading us to finalize our 6-factor model with Sigma-3, B/Mm, CEI, Log_MV, Alpha_3, and RSI as the selected explanatory variables.

In addition, we took into account that size and momentum showed high significance

[54] See Fama and French (2015).

in the double sort procedure as referenced in Table 2. Consequently, we constructed an alternative three-factor model that incorporated size and momentum-related variables, i.e., Log_MV, Alpha_3, and RSI. It enabled us to examine the performance of a more compact set of variables. Consequently, both the comprehensive six-factor model and the more compact three-factor model will be employed in our subsequent analysis.

3.3 Stock Selection

We used Lewellen's method and VAR method to select the stock.

3.3.1 Lewellen's Regression

In this chapter, we explained how to forecast expected returns using Lewellen's method. We used the method of forecasting expected returns using Lewellen (2015). The model uses the 6 variables (B/Mm, Log_MV, CEI,Sigma_3, Alpha_3, RSI) and 3 variables (Log_MV, Alpha_3, RSI) chosen in factor selection.

We excluded microcap stocks from the dataset. So, we used 97% of the sum of the capitalization of stocks in the United Kingdom.

First, we regressed excess return on characteristics in the cross-section following Fama and MacBeth (1973)[55]. Then we obtained γ, the same meaning as the coefficient in the fama-Macbeth regression model for each month.

$$\begin{gathered} R_{i,1} = \gamma_{1,0} + \gamma_{1,1} BMm_0 + \gamma_{1,2} logMV_0 + \cdots + \varepsilon_{i,1} \\ R_{i,2} = \gamma_{2,0} + \gamma_{2,1} BMm_1 + \gamma_{2,2} logMV_1 + \cdots + \varepsilon_{i,2} \\ \vdots \\ R_{i,T} = \gamma_{T,0} + \gamma_{T,1} BMm_{T-1} + \gamma_{T,2} logMV_{T-1} + \cdots + \varepsilon_{i,1} \end{gathered} \tag{6}$$

Next, we computed the average of coefficients. Here we used two differences. The study by Lewellen (2015) uses a rolling average and a cumulative average[56]. Therefore we decided to use the same calculation method. However, unlike Lewellen, we used 36 months for the rolling average period. This is based on the Hanuer and Lauterbach(2019) criterion[57], which allows the model to be flexible in forecasting over a shorter time.

$$\widehat{\gamma_0}_{cumulative} = \frac{1}{36} \sum_{t=n-35}^{n} \gamma_{t,0}, \widehat{\gamma_1} = \frac{1}{T} \sum_{t=n-35}^{n} \gamma_{t,1}, \cdots \widehat{\gamma_m} = \frac{1}{T} \sum_{t=n-35}^{n} \gamma_{t,m}, \tag{7}$$

$$\widehat{\gamma_0}_{rolling} = \frac{1}{T} \sum_{t=1}^{T} \gamma_{t,0}, \widehat{\gamma_1} = \frac{1}{T} \sum_{t=1}^{T} \gamma_{t,1}, \cdots \widehat{\gamma_m} = \frac{1}{T} \sum_{t=1}^{T} \gamma_{t,m}, \tag{8}$$

55 See Fama and MacBeth (1973).
56 See Lewellen (2015).
57 See Hanauer and Lauterbach (2019).

	Risk			Value				Profitability			Investment		Size		Momentum		
Model	Beta_3	Sigma_3	SD	B/M	B/Mm	E/Pm	C/Pm	OP/BE	NOA	ROE	AG	CEI	MV	Log_MV	Mom_12_2	Alpha_3	RSI
Panel A: Five-Factor Model Variables																	
(1)	5.55E-01			1.14E-01				1.22E-01			-3.75E-03		2.28E-05				
	7.91E-02			3.00E-01				3.70E-01			1.02E-01		7.59E-02				
Panel B: Six-Factor Model Variables																	
(2)	4.96E-01			1.54E-01				1.73E-01			-2.73E-03		2.10E-05		1.09E-02		
	1.27E-01			1.44E-01				2.10E-01			2.38E-01		9.21E-02		7.54E-08		
Panel C: Comparison Risk Variables																	
(3)	5.49E-01	1.00E+00	1.85E-02	-4.51E-01				1.44E-01			2.03E-03		-6.27E-06		2.39E-02		
	2.59E-02	< <2.2e-16	6.18E-08	1.41E-81				1.15E-06			1.68E-05		1.76E-04		3.70E-191		
Panel D: Comparison Value Variables																	
(4)	5.50E-01	5.50E-01	2.26E-02	3.23E-01	-1.00E+00			1.17E-01			1.83E-03		-5.16E-06		2.11E-02		
	2.44E-02	2.44E-02	5.85E-11	3.47E-15	1.86E-76			1.50E-05			1.28E-04		1.67E-03		6.60E-158		
(5)	-6.82E-01	9.34E-01	1.61E-01		-1.12E+00	3.58E+00	1.58E+01	-1.37E+00			1.80E-02		-1.52E-04		3.02E-02		
	4.08E-01	1.20E-31	1.37E-01		1.67E-03	9.52E-02	2.74E-01	2.15E-01			1.94E-01		2.08E-01		6.04E-03		
Panel E: Comparison Profitability Variables																	
(6)		9.87E-01	5.22E-02		-8.05E-01			-1.28E-01	5.05E-08	3.30E-01	9.50E-04		-3.42E-05		2.02E-02		
		< <2.2e-16	2.71E-05		3.82E-72			9.38E-03	5.12E-06	1.22E-04	2.55E-01		1.03E-08		1.05E-78		
Panel F: Comparison Investment Variables																	
(7)		9.91E-01	4.05E-02		-7.18E-01			-1.07E-01	4.01E-08	2.63E-01	1.36E-03	-1.08E-01	-2.22E-05		1.71E-02		
		< <2.2e-16	6.04E-04		5.94E-69			2.17E-02	1.87E-04	2.36E-03	1.00E-01	3.64E-44	1.13E-04		2.51E-52		
Panel G: Comparison Size Variables																	
(8)		9.90E-01	4.03E-02		-7.07E-01			-8.34E-02	2.40E-08	2.71E-01		-1.09E-01	2.32E-06	-5.63E-02	1.71E-02		
		< <2.2e-16	8.92E-04		6.10E-66			9.88E-02	4.63E-02	2.05E-03		1.20E-46	8.15E-01	1.08E-03	3.81E-52		
Panel H: Comparison Momentum Variables																	
(9)		9.86E-01	-1.32E-02		-2.19E-01				1.05E-08	3.14E-02		-1.80E-02		-4.50E-02	-6.39E-04	6.64E-01	2.32E-02
		< <2.2e-16	8.67E-02		1.14E-13				4.24E-06	3.56E-01		2.56E-04		3.35E-09	5.65E-01	4.80E-66	1.05E-24
Panel I: Strongest EM Variables																	
(10)		9.86E-01			-1.90E-01				1.15E-08			-1.62E-02		-4.91E-02		6.80E-01	2.07E-02
		< <2.2e-16			2.77E-09				4.74E-08			1.04E-02		1.19E-07		3.85E-60	6.40E-17
Panel J: Only big stocks																	
(11)	5.61E-01			5.28E-02				2.37E-02			9.01E-05		1.67E-05				
	9.73E-02			6.97E-01				9.02E-01			9.77E-01		1.59E-01				
(12)		9.84E-01			-1.89E-01				8.47E-09			-1.96E-02		-2.82E-02		6.87E-01	1.83E-02
		< <2.2e-16			2.29E-10				3.88E-06			1.73E-03		4.47E-02		3.76E-55	9.51E-13

Table 3: Results of the cross-sectional regressions

where m denotes the number of factors. Next, the calculated average coefficient is multiplied by the latest firm characteristics to calculate the expected return. This replicates the portfolio manager, who can use only previous data then.

$$\widehat{R_{l,T}} = \gamma_{\widehat{T-1},0} + \gamma_{\widehat{T-1},1} \times B/Mm_{T-1} + \gamma_{\widehat{T-1},2} \times logMV_{T-1} + \cdots + \varepsilon_{i,1} \tag{9}$$

The critical point is that the accuracy of our model tells us which stocks will have higher expected returns in a current month.

Finally, we sorted the expected returns and determined the top 20 stocks for stock selection.

3.3.2 Vector Autoregression

Vector autoregression (VAR) is a valuable method utilized for predicting the dynamics of an interconnected system using historical time series data. In our context, we have six explanatory factors, each correlated with the return variable. Although we have not formally tested the intercorrelation among these factors, we can still leverage VAR to forecast these variables, particularly the return.

As we discussed earlier, we initially considered both a 3-factor model and a 6-factor model. However, we encountered a notable number of missing values in the B/Mm factor. Since VAR heavily relies on time series data and can be sensitive to missing values, we have judiciously chosen to omit the B/Mm factor from the 6-factor model. Consequently, we now employ a 5-factor model to ensure data completeness and robustness in our analysis.

The 3-factor VAR model is defined as:

$$\begin{aligned}\begin{bmatrix} R_t \\ LogMV_t \\ Alpha3_t \\ RSI_t \end{bmatrix} &= C + A_1 \begin{bmatrix} R_{t-1} \\ LogMV_{t-1} \\ Alpha3_{t-1} \\ RSI_{t-1} \end{bmatrix} + A_2 \begin{bmatrix} R_{t-2} \\ LogMV_{t-2} \\ Alpha3_{t-2} \\ RSI_{t-2} \end{bmatrix} + \cdots \\ &+ A_{p-1} \begin{bmatrix} R_{t-p+1} \\ logMV_{t-p+1} \\ Alpha3_{t-p+1} \\ RSI_{t-p+1} \end{bmatrix} + A_p \begin{bmatrix} R_{t-p} \\ logMV_{t-p} \\ Alpha3_{t-p} \\ RSI_{t-p} \end{bmatrix} + e_t.\end{aligned} \tag{10}$$

The 5-factor VAR model is defined as:

$$
\begin{aligned}
\begin{bmatrix} R_t \\ Sigma3_t \\ CEI_t \\ LogMV_t \\ Alpha3_t \\ RSI_t \end{bmatrix} &= C + A_1 \begin{bmatrix} R_{t-1} \\ Sigma_{t-1} \\ CEI_{t-1} \\ LogMV_{t-1} \\ Alpha3_{t-1} \\ RSI_{t-1} \end{bmatrix} + A_2 \begin{bmatrix} R_{t-2} \\ Sigma_{t-2} \\ CEI_{t-2} \\ LogMV_{t-2} \\ Alpha3_{t-2} \\ RSI_{t-2} \end{bmatrix} + \cdots \\
&+ A_{p-1} \begin{bmatrix} R_{t-p+1} \\ Sigma_{t-p+1} \\ CEI_{t-p+1} \\ LogMV_{t-p+1} \\ Alpha3_{t-p+1} \\ FSI \end{bmatrix} + A_p \begin{bmatrix} R_{t-p} \\ Sigma_{t-p} \\ CEI_{t-p} \\ LogMV_{t-p} \\ Alpha3_{t-p} \\ FSI \end{bmatrix} + e_t.
\end{aligned}
\tag{11}
$$

For each stock, we applied these two methods to predict the return. The process can be described as follows. Firstly, we used the first 36-month data to fit model and predict the value in the 37th month. Then, we used the first (n-1) month data to fit the model and predict the return in the nth month. Finally, we obtained a panel data of return for each stock and each time points. For each time point, we selected the top 20 stocks and put them into the portfolio.

In our research, we applied these two methods to predict the return for each individual stock. The step-by-step process is outlined as follows:

1. We started by utilizing the data from the initial 36 months to train the predictive model. Using this model, we made a forecast for the return value in the 37th month.

2. Then, we adopted a cumulative window approach, where we used data from the first (n-1) months to fit the model and made a return prediction for the nth month. By repeating this process for various time points and making short-term prediction, we enhanced the model's consistency and effectiveness in making predictions.

3. By applying the above methods for each individual stock, we assembled a comprehensive panel dataset. This dataset encompasses predicted returns for each stock across various time points, forming the basis for our analysis.

4. At each time point, we selected the top 20 performing stocks based on the model's predictions. These selected stocks were then pooled together to form a portfolio.

This portfolio construction enabled us to focus on stocks with the most promising return potential, with the aim of achieving superior portfolio performance.

3.4 Portfolio Further Optimization

Every month we selected top 20 stocks by iteratively implementing Lewellen's method and Vector Autoregression Model. In other words, stocks in portfolio would be updated every month, which allowed us to achieve a dynamic portfolio composition that stayed in line with our investment objectives.

To optimize our investment portfolio further, we adopted the Tangent Portfolio Method. This technique enabled us to find the perfect balance between risk and return, maximizing the risk-return trade-off. The Tangent Portfolio Method involved drawing a line tangent to the efficient frontier, representing the set of optimal portfolios that offer the highest return for a given level of risk.

Through the Tangent Portfolio Method, we constructed portfolios that consistently aligned with our desired risk-return trade-off. This approach allowed us to strike a balance between maximizing returns and minimizing risks, essential for achieving long-term investment success.

As investment portfolios are exposed to market fluctuations, the risk and return characteristics of assets within the portfolio can change over time. In order to adapt to market volatility and limit the portfolio turnover, we quarterly rebalanced the portfolio to ensure our product remain aligned with best risk-return strategy without generating too much transaction cost.

To construct our portfolio, we use the past 24 months' returns, volatility and risk-free rates as inputs for the Tangent Portfolio Method. However, during specific periods, as the Figure 1 shows, such as financial crises, negative expected returns and non-shorting requirement may make it infeasible to calculate the tangent portfolio. The mean of minimum variance portfolio return is negative. In such cases, the efficient frontier may fall below the x-axis, making it impossible to derive a tangent portfolio. During these periods, we maintained the stock portfolio without rebalancing and reweighting. By employing these portfolio optimization techniques mentioned above, 6 portfolio candidates were constructed. These portfolios are now ready to undergo simulation to assess their historical performance over the period from 1990 to 2021.

The subsequent chapter will present empirical evidence of our investment strategy, offering insights into the actual performance of each portfolio candidates. Through rigorous analysis and evaluation, we will identify and select the most effective strategy among the six candidates.

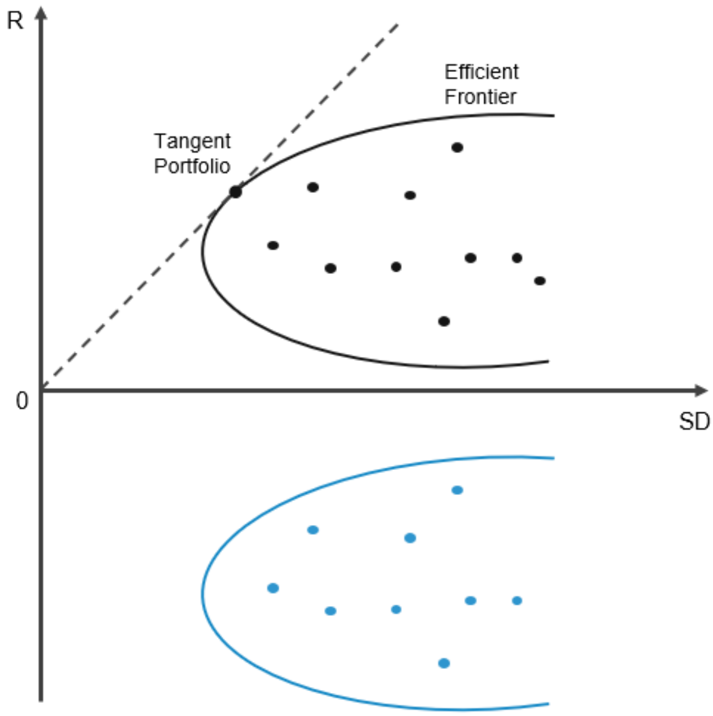

Figure 1: Tangent portfolio illustration

4 Empirical Evidence

4.1 Visualize Six Strategies' Cumulative Performance

4.1.1 Six Strategy Portfolios Summary

Previous steps yielded us six strategy candidates. The six strategies and corresponding abbreviation are presented below in Table 4. Columns represent factors and rows represent two forecasting strategies.

	3–Factor	5–Factor	6–Factor
Vector Autoregression	X3_var	X5_var	
Lewellen's Method (Rolling)	X3_roll		X6_roll
Lewellen's Method (Cumulative)	X3_cum		X6_cum

Table 4: Six strategies overview

Note: 3-factor denotes log_MV, Alpha_3, RSI, 5-factor denotes Sigma_3, CEI, log_MV, Alpha_3, RSI, and 6-factor denotes Sigma_3, B/Mm, CEI, log_MV, Alpha_3, RSI.

4.1.2 Cumulative Performance Visualization

We sought to evaluate the effectiveness and potential benefits of six distinct strategies within our portfolio. To achieve this, we conducted a simulation spanning the period from January 1990 to January 2021, during which we accumulated the return and tracked the evolution of each strategy.

Figure 2 visualized that the cumulative returns(%) of our portfolio, from where we have some observations. Firstly, the majority of our strategies outperformed the benchmark portfolio, demonstrating their potential to generate superior returns. However, it's worth noting that the 3-Factor VAR Strategy (X3_var) did not perform as well as the benchmark. Of particular significance were the exceptional per-

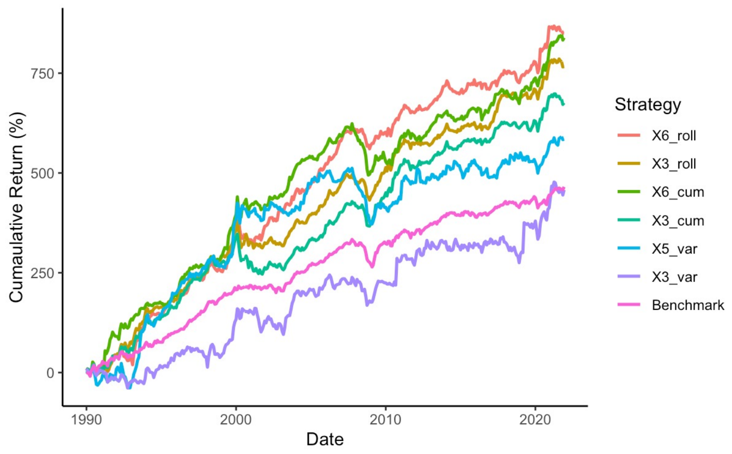

Figure 2: Six strategies cumulative return visualization

Note: Here we select big stocks which together account for 90% of a country's aggregated market capitalization as benchmark portfolio, so the number of stocks in the benchmark is different from our strategy.

formances of two strategies: the 6-Factor Rolling-Window Lewellen's Regression Strategy (X6_roll) and the 6-Factor Cumulative Lewellen's Regression Strategy (X6_cum). These two strategies exhibited significant outperformance compared to the other strategies in the portfolio. However, it was evident that the 6-Factor Rolling-Window Lewellen's Regression Strategy (X6_roll) experienced a drastic plummet during the 2008 financial crisis, in contrast to the more resilient performance of the 6-Factor Cumulative Lewellen's Regression Strategy (X6_cum), which maintained its upward momentum with minimal drawdowns.

After financial crisis, the 6-Factor Rolling-Window Lewellen's Regression Strategy (X6_roll) emerged as the most favourable strategy, generating the highest returns among all the strategies in the portfolio.

Additionally, we assessed the impact of different weighting and rebalancing strategies. To compare the alternative weighting and rebalancing strategy, we controlled stock selection strategy and constructed the 6-Factor Cumulative Lewellen's Regression Strategy (X6_cum) with equally weighted strategy. Figure 3 revealed that while equally weighted portfolios offer better diversification and risk reduction, in our specific case, the strategy that maximized the Sharpe ratio proved to be the most effective. Without the application of the Sharpe ratio, our portfolio failed to beat the benchmark, highlighting its crucial role in enhancing overall performance.

Overall, in the UK market context, strategies incorporating a greater number of factors exhibit superior performance compared to those with a lower number of factors. Moreover, strategies employing Lewellen's method demonstrate stronger predictive capabilities when contrasted with those utilizing Vector Autoregression.

Figure 3: Equally weighted and maximum sharpe ratio weighted strategies

4.1.3 Portfolio Characteristic

Performance	X6_roll	X6_cum	X3_roll	X3_cum	X5_var	X3_var	Benchmark
Average Number of Stock	5.27	5.27	7.86	7.96	8.26	7.68	215
Ann. Turnover	0.91	0.93	0.94	0.91	0.83	0.84	0.11
Effective N	3.64	3.99	5.06	5.39	5.94	4.65	50.12

Table 5: Portfolio strategies vs benchmark characteristic

Table 5 shows the characteristic of six portfolio strategies and benchmark. The column displays portfolio strategies. The average number of stocks are quite small and we strongly focus on several specific stocks. The annualized turnover measures the average one-way portfolio turnover observed for each strategy. Lower turnover rates generally indicate lower trading activity and fewer changes in portfolio holdings. Here we calculated the annualized turnover rate according to DeMiguel et al. (2009)[58]:

$$Turnover_i := \frac{1}{2}\frac{1}{T}\sum_{t=1}^{T}\sum_{n=1}^{N}(|w_{n,t+1,i} - w_{n,t,i}|) \tag{12}$$

Effective N is the average reciprocal of the Hirshman-Herfindahl index of portfolio weights. The lower effective N indicates the higher concentration on certain stocks with a portfolio. Here we calculated the annualized turnover rate according to Chow et al. (2016)[59]:

$$EffectiveN_i := \frac{1}{T}\sum_{t=1}^{T}\left(\sum_{n=1}^{N} w_{n,t,i}^2\right)^{-1} \tag{13}$$

[58] See Demiguel et al. (2009).
[59] See Chow et al. (2016).

4.2 Six Strategies Performance Statistic

Table 6 shows arithmetic returns, standard deviations, Sharp ratio, Outperformance, tracking error, information ratio, and VaR for six strategies and benchmarks. All figures are calculated based on annualized return data. The sharp ratio is calculated by dividing return by volatility. Outperformance is the average difference between portfolio and market returns (”active return”), which means how much return we earned above market performance. The standard deviation of active return calculates tracking error, indicating how much a portfolio’s returns move differently from the market. The Information Ratio is calculated by dividing the Outperformance by the Tracking error. It means how much more return is earned relative to the risk considered in the tracking error. VaR is calculated in 95% of confidence level, which means the maximum expected loss of portfolio value.

Performance	X6_roll	X6_cum	X3_roll	X3_cum	X5_var	X3_var	Benchmark
Return (%)	32.65	31.91	28.86	25.64	24.68	14.44	15.82
Volatility (%)	45.26	43.56	38.66	37.36	66.39	38.91	19.20
SR	0.72	0.73	0.74	0.68	0.37	0.37	0.82
Outperformance (%)	16.82	16.09	13.04	9.82	8.86	-13.76	
TE (%)	39.52	35.88	29.60	27.83	62.40	32.94	
IR	0.42	0.44	0.44	0.32	0.14	-0.041	
VaR (%)	-31.68	-18.78	-22.59	-29.99	-33.58	-34.43	-7.70

Table 6: Performance of our strategies and benchmarks

The table presents performance and risk metrics for the 6 investment strategies and benchmark tested in out study. Return (%) is the average arithmetic return (annualized) as a percentage. Volatility (%) is the standard deviation (annualized) as a percentage. SR is the annualized Sharp Ratio. Outperformance (%) indicates the difference of portfolio return and market return which is ”active return”. TE (%) represents the tracking error to the value-weighted portfolio portfolio as a percentage. IR is the information ratio with respect value-weighted portfolio.

$$TE = \sigma\left(R_p - R_B\right) \tag{14}$$

$$IR = \frac{\overline{R_p} - \overline{R_B}}{TE} \tag{15}$$

The results shows that X6_roll strategy has the highest average return (32.65%) while X3_cum has minimum volatility in 6 strategies. X3_mean stands out with the highest Sharpe Ratio of 0.74 among all strategies, showcasing its superior risk-adjusted performance. X3_var has the smallest average return and the largest VaR (-34.43%), implying a higher maximum expected loss.

Overall, our strategy outperforms the benchmark in return. However, the volatility is higher than the benchmark. Therefore, the Sharpe ratio, a comprehensive measure of return and risk, did not outperform the benchmark. The average outperformance of the Lewellen strategy compared to the benchmark is around 9%. Tracking error is also high, meaning six strategies move differently from the market return. We believed that the average number of stocks is one of the reasons for this difference from the benchmark. All strategies have between 5 and 8 stocks, which is a tiny number. Therefore, our portfolios have aggressive behavior when compared to the benchmark. Based on the Cumulative Return and Average annualized return in the time series, we assumed that the "6-factor rolling window" was the best strategy.

4.3 Style Analysis

In 1989, Sharpe[60] pioneered the returns-based style analysis as a method to evaluate the behavior of portfolio managers. Style analysis is a statistical approach used to attribute the returns of a portfolio to different explanatory factors. Through this technique, we derived the portfolio's exposures to these factors, providing insights into the investment style adopted by the portfolio. In our specific case, we performed a regression analysis by regressing the portfolio's return on the Fama-French five factors (market risk, size, value, profitability, and investment) and momentum. This regression equation takes the following form:

$$R_t = \alpha + \beta_1 RMRF_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 CMA_t + \beta_5 RMW_t + \beta_6 WML_t, \quad (16)$$

where α represents the intercept, indicating the portfolio's excess return not explained by the selected factors and β_1, ..., β_6 denote the regression coefficients, reflecting the portfolio's exposures to each respective factor.

The regression result is suggested in Table 7. It reveals compelling results indicating the significance of specific factors in influencing our portfolio's return. Specifically, the factors RMRF, SMB, and WML exhibit significantly positive effects, while HML demonstrates a significant negative impact on the portfolio's return. These findings imply that our portfolio benefits from exposure to market risk premium, small-cap stocks, low-value stocks, and stocks with a momentum effect.

Besides, an important observation from the analysis is the presence of a positively significant abnormal return alpha. It suggests that there are unaccounted-for factors that positively impact the portfolio's returns, beyond the influence of RMRF, SMB, HML, and WML.

5 Conclusion and Future Work

This study embarked on an exploration of factor-based investment strategies in the context of the UK stock market. Our primary aim was to create an investment strategy that maximizes returns while efficiently managing risk.

We adopted a dynamic portfolio composition approach, where portfolios were updated monthly using Lewellen's method and vector autoregression model. We then

[60] See Sharpe (1988).

	Coefficient Estimate	Standard Error	t value	P value
Intercept	2.82	0.39	7.25	<2.32e-12
RMRF	0.82	0.07	11.53	<2e-16
SMB	0.51	0.15	3.38	<0.000795
HML	-0.65	0.14	-4.58	<6.40e-06
CMA	-0.13	0.17	-0.79	0.429052
RMW	-0.32	0.18	-1.84	0.066565
WML	0.20	0.09	2.36	<0.018891
Multiple R-squared: 0.342				
Adjusted R-squared: 0.3315				

Table 7: Style analysis

employed the tangent portfolio method to optimize the portfolios and rebalanced them quarterly. The portfolios were analyzed using a variety of performance metrics over the period of 1990 to 2021.

Our empirical results indicated that the majority of our strategies outperformed the benchmark portfolio, highlighting the effectiveness of our selected strategies. Of particular note were the 6-Factor Rolling-Window Lewellen's Regression Strategy (X6_roll) and the 6-Factor Cumulative Lewellen's Regression Strategy (X6_cum), both of which exhibited superior performance.

Furthermore, we found that strategies incorporating a greater number of factors tend to perform better than those with fewer factors in the context of the UK stock market. The superior predictive abilities of strategies based on Lewellen's method compared to those using vector autoregression model were also noteworthy.

While our investment strategy outperformed the benchmark in terms of return, it did bear a higher degree of volatility, which is a potential area for improvement. Lastly, the style analysis indicated that our portfolio benefits from exposure to market risk premium, low-value stocks, and stocks with a momentum effect.

This research opens several paths for future work. While we found significant success with our investment strategies, future research could explore ways to further mitigate volatility without hampering returns. This could involve considering other risk factors or diversification strategies based on investment preferences not specified in this study. Future research could also involve a deeper exploration into the impact of different weighting and rebalancing strategies.

Additionally, it would be beneficial to compare our strategies' performance in other market contexts beyond the UK. It would provide insights into the adaptability and robustness of the strategies under varying market conditions and structures.

It should be emphasized that if more firm characteristics are available in the dataset, the implementation of advanced techniques, such as machine learning methods, is also possible to be incorporated into the portfolio construction process. Otherwise, the projection of stock returns could be poor.[61]

Lastly, the impact of other potential factors on portfolio returns could be explored further. As this study has shown, certain factors have significant impacts on portfolio returns. Future work could delve into identifying these unaccounted-for factors that could boost portfolio performance.

By continuing to explore these avenues, we can gain a deeper understanding of factor-based investment strategies and their practical implications, and further improve the design of investment strategies.

[61] See Rasekhschaffe and Jones (2019); Gu et al. (2020).

Appendix: Variable Definitions

Appendix 1: Variable Definitions

Table 8 shows the definition of variables used in our work. Furthermore, below we provide more information regarding the underlying data used to calculate the above mentioned variables.

Value

Book Equity is defined as common equity (WC03501) plus deferred taxes (WC03263).

Momentum

RSI is Relative Strength Index which is a momentum oscillator developed in Wilder (1978)[62].

RS is average positive return last one year divided by average negative return last one year.

Quality

Gross profit is defined as net sales or revenues (WC01001) minus cost of goods sold(WC01501).

We measure operating profits as sales or revenues (WC01001) minus cost of goods sold (WC01051), minus selling, general, and administrative expenses (WC01101), minus interest expense (WC01251).

Operating assets is total assets (WC02999) minus cash and short-term investment (WC02001). Operating liabilities is total assets minus short-term and long- term debt (WC03255), minus minority interest (WC03426), minus preferred stock and common equity (WC03995).

We define Change in operating working capital is the change in current assets (WC02201) minus change in cash and short-term investments (WC02001), minus change in current liabilities (WC03101), plus change in debt in current liabilities (WC03051), plus change in income taxes payable (WC03063, zero if missing).

Investment

We define Split-adjusted shares outstanding as shares outstanding (Datastream item NOSH) divided by the adjustment factor (Datastream item AF). We measure change in gross property, plant, and equipment (WC02301) plus the annual change in inventories (WC02101) (both from fiscal year ending in calendar year $y-2$ to

62 See Wilder (1978).

fiscal year ending in calendar year y − 1) all divided by total assets (WC02999) of year y − 2.

MSCI Category	Abbreviation	description	calculation or source	source
Value	B/M	Book to market ratio	$\frac{\text{Book Equity}}{\text{MV in last December}}$	B
	B/Mm	Book to market ratio using the most recent MV	$\frac{\text{Book Equity}}{\text{MV}}$	H
	E/P	Earnings to price	$\frac{\text{Earnigns before extraordinary items(WC01551)}}{\text{MV in last December}}$	B
	E/Pm	Earnings to price using the most recent MV	$\frac{\text{Earnigns before extraordinary items(WC01551)}}{\text{MV}}$	H
	C/P	Cash flow-to-price	$\frac{\text{Operating cash flow(WC04860)}}{\text{MV in last December}}$	H
	C/Pm	Cash flow-to-price using the most recent MV	$\frac{\text{Operating cash flow(WC04860)}}{\text{MV}}$	H
	EBITDA/EV	Enterprise Multiple	$\frac{\text{EBITDA(WC18198)}}{\text{EV(WC18100)}}$	M
Size	Log_MV	Log of market value	$\log(\text{MV})$	M
	MV	Market value	MV	H
Momentum	RSI	relevant strength index	Momentum oscillator determining the pace and variation of stock price. $100 - \frac{100}{1+\text{RS}}$	M
	Alpha_3	Historical alpha (3 year)	Values of historical alpha which are computed by the time-series regression.	M
	Mom_12_2	Momentum_12_2	Geometric mean of the compound return of a stock over the past twelve months, but ignores the previous month.	H
Yield	DY	Dividend yield	$\frac{\text{Dividends Per Share(WC05101)}}{\text{Price Per Share(WC05001)}}$	M
Quality	ROE	Return on equity	$\frac{\text{Earnigns before extraordinary items(WC01551)}}{\text{Book Equity}}$	B
	ROA	Return on asset	$\frac{\text{Earnigns before extraordinary items(WC01551)}}{\text{Total Asset(WC02999)}}$	B
	GP/A	Gross profit to asset Novy-Marx (2013)	$\frac{\text{Gross profit}}{\text{Total Asset(WC02999)}}$	H
	OP/BE	Operating profit-to-book equity	$\frac{\text{Operating profits}}{\text{Book Equity}}$	H
	NOA	Net operating asset	Operating Asset $-$ Operating Liability	H
	OA	Operating accrual	ΔOperating Working Capital $-$ Lagged Depreciation	H

Continued on next page

Table 8: Variable definitions in univariate sort.

MSCI Category	Abbreviation	description	calculation or source	source
Volatility	Beta_3	Beta (3 year)	Slope coefficient from a time-series regression of stock excess returns, against the value-weighted excess returns	B
	Sigma_3	Sigma (3 year)	Volatility of the residual returns from historical beta equation	M
	SD	Monthly standard deviation	Volatility of excess returns over past year	B
Growth	EPS	Historical earnings per share growth rate	$\frac{\text{EPS(WC05202)}-\text{EPS in Last Year}}{\text{EPS in Last Year}}$	M
	SPS	Historical sales per share growth rate	$\frac{\text{SPS(WC05508)}-\text{SPS in Last Year}}{\text{SPS in Last Year}}$	M
Investment	AG	Asset growth	$\frac{\text{Total Asset}_{y-1}-\text{Total Asset}_{y-2}}{\text{Total Asset}_{y-2}}$	H
	NSI	Net stock issues	$\log(\frac{\text{Sprit Share}}{\text{Sprit Share Last Year}})$	H
	CEI	Composite equity issuance	$\log(\text{MV}_y - \text{MV}_{y-1}) - R_{y-1,y}$	H
	I/A	Investment to assets	$\frac{\triangle\text{Gross property}+\text{Annual change in inventory}}{\text{Total Asset}_{y-1}}$	H

Table 8: Variable definitions in univariate sort (Continued)

Note: M denotes Variable comes from MSCI factor framework, H denotes variable comes from Hanuer and Lauterbach's paper, and B denotes variable used in both Hanauer and Lauterbach's paper and MSCI.

References

Antonelli, S., Corneli, F., Ferriani, F., & Gazzani, A. (2022). Benchmark effects from the inclusion of chinese a-shares in the msci em index. *Economics Letters*, *216*, 110600.

Asness, C. S., & Frazzini, A. (2013). The devil in hml's details. *The Journal of Portfolio Management*, *39*, 49–68.

Başçı, E. S., & Karaca, S. S. (2013). The determinants of stock market index : Var approach to turkish stock market. *International journal of economics and financial issues : IJEFI*, *3*(1).

Bender, J., Briand, R., Melas, D., & Subramanian, R. A. (2013). Foundations of factor investing. *SSRN Electronic Journal*.

Bisnoff, J. (2022). *Robinhood losses spell the end of an era for young investors who've never traded through a downturn*. https://www.forbes.com/sites/jasonbisnoff/2022/04/29/robinhood-losses-spell-the-end-of-an-era-for-young-investors-whove-never-traded-through-a-downturn/?sh=629c37dc5bea

Blackburn, D. W., & Cakici, N. (2017). Overreaction and the cross-section of returns: International evidence. *Journal of Empirical Finance*, *42*, 1–14.

Board of Governors of the Federal Reserve System (US). (2023). *6-month treasury bill secondary market rate, discount basis [dtb6], retrieved from fred, federal reserve bank of st. louis*. https://fred.stlouisfed.org/series/DTB6

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

Choueifaty, Y., & Coignard, Y. (2008). Toward maximum diversification. *The Journal of Portfolio Management*, *35*(1), 40–51.

Chow, T.-M., Kose, E., & Li, F. (2016). The impact of constraints on minimum-variance portfolios. *Financial Analysts Journal*, *72*(2), 52–70.

Cooper, M. J., Gulen, H., & Schill, M. J. (2008). Asset growth and the cross-section of stock returns. *The Journal of Finance*, *63*(4), 1609–1651. Retrieved July 24, 2023, from http://www.jstor.org/stable/25094485

Daniel, K., & Titman, S. (2006). Market reactions to tangible and intangible information. *The Journal of Finance*, *61*(4), 1605–1643. Retrieved July 24, 2023, from http://www.jstor.org/stable/3874722

Demiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *Review of Financial Studies*, *22*.

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

Fama, E. F., & French, K. R. (2008). Dissecting anomalies. *The Journal of Finance*, *63*(4), 1653–1678. Retrieved July 24, 2023, from http://www.jstor.org/stable/25094486

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

Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*, *81*(3), 607–636. Retrieved July 22, 2023, from http://www.jstor.org/stable/1831028

Friend, I., & Blume, M. E. (1970). Measurement of portfolio performance under uncertainty. *American Economic Review*, *60*(4), 561–75.

Gebhardt, W. R., Hvidkjaer, S., & Swaminathan, B. (2005). The cross-section of expected corporate bond returns: Betas or characteristics? *Journal of Financial Economics*, *75*(1), 85–114.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, *33*(5), 2223–2273.

Hanauer, M. X., & Lauterbach, J. G. (2019). The cross-section of emerging market stock returns. *Emerging Markets Review*, *38*, 265–286.

Hirshleifer, D., Hou, K., Teoh, S. H., & Zhang, Y. (2004). Do investors overvalue firms with bloated balance sheets? *Journal of Accounting and Economics*, *38*(1), 297–331. https://EconPapers.repec.org/RePEc:eee:jaecon:v:38:y:2004:i::p:297-331

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 July 24, 2023, from http://www.jstor.org/stable/2328882

Jensen, M. C., Black, F., & Scholes, M. S. (2006). The capital asset pricing model: Some empirical tests. *Capital Markets: Asset Pricing & Valuation*.

Jung, W.-S., Chae, S., Yang, J.-S., & Moon, H.-T. (2006). Characteristics of the korean stock market correlations. *Physica A: Statistical Mechanics and its Applications*, *361*(1), 263–271.

Lewellen, J. (2015). The cross-section of expected stock returns. *Critical Finance Review*, *4*(1), 1–44.

McLean, R. D., & Pontiff, J. (2016). Does academic research destroy stock return predictability? *The Journal of Finance*, *71*(1), 5–32.

Meek, G. K., & Gray, S. J. (1989). Globalization of stock markets and foreign listing requirements: Voluntary disclosures by continental european companies listed on the london stock exchange. *Journal of International Business Studies*, *20*(2), 315–336.

MSCI Inc. (2023). *Factor Investing*. https://www.msci.com/our-solutions/factor-investing

Nartea, G. V., Kong, D., & Wu, J. (2017). Do extreme returns matter in emerging markets? evidence from the chinese stock market. *Journal of Banking & Finance*, *76*, 189–197.

Ogryczak, W., Przyluski, M., & Sliwinski, T. (2015). Portfolio optimization with reward-risk ratio measure based on the conditional value-at-risk.

Pasquariello, P. (2000). The fama-macbeth approach revisited.

Pontiff, J., & Woodgate, A. (2008). Share issuance and cross-sectional returns. *The Journal of Finance*, *63*(2), 921–945.

Rapach, D., & Zhou, G. (2023). Forecasting stock returns. *SSRN Electronic Journal*.

Rasekhschaffe, K. C., & Jones, R. C. (2019). Machine learning for stock selection. *Financial Analysts Journal*, *75*(3), 70–88.

Sharpe, W. F. (1988). Determining a fund's effective asset mix. *Investment management review*, *2*(6), 59–69.

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, *71*(3), 289–315. Retrieved July 24, 2023, from http://www.jstor.org/stable/248290

Slotte, P. (2021). Asset growth anomaly in the uk stock market.

Sullivan, E. J. (2007). A brief history of the capital asset pricing model.

Theron, L., & van Vuuren, G. (2018). The maximum diversification investment strategy: A portfolio performance comparison (D. McMillan, Ed.). *Cogent Economics &amp Finance*, *6*(1), 1427533.

Wang, C., & Low, S. S. (2003). Hedging with foreign currency denominated stock index futures: Evidence from the MSCI taiwan index futures market. *Journal of Multinational Financial Management*, *13*(1), 1–17.

Wilder, J. W. (1978). New concepts in technical trading systems. *(No Title)*.

Yang, J. (2003). Market segmentation and information asymmetry in chinese stock markets: A VAR analysis. *SSRN Electronic Journal*.

Zivot, E., & Wang, J. (2006). Vector autoregressive models for multivariate time series. In *Modeling financial time series with s-PLUS®* (pp. 385–429). Springer New York.

Declaration of Academic Integrity

Hereby, I declare that I have composed the presented paper independently on my own and without any other resources than the ones indicated. All thoughts taken directly or indirectly from external sources are properly denoted as such.

This paper has neither been previously submitted to another authority nor has it been published yet.

Munich, July 20, 2023 Signature: CHEN Yuru

Munich, July 20, 2023 Signature: [illegible]

Munich, July 20, 2023 Signature: Yasuaki

Munich, July 20, 2023 Signature: [illegible]

Munich, July 20, 2023 Signature: [illegible]