Factor Momentum and Mean Reversion in Multifactor Investing*

Project Studies

Niklas Baier
TUM School of Management
Technical University of Munich
Munich, Germany
niklas.baier@tum.de

Tobias Baldauf
TUM School of Management
Technical University of Munich
Munich, Germany
tobias.baldauf@tum.de

April 10, 2022

Abstract

This research report explores potential mean reversion effects in factor momentum investing. On a sample of 10 factors built from the S&P 500 universe over the past 20 years, we define base and maximum Sharpe ratio variants of cross-sectional and time-series factor momentum which are then used to construct mean reversion strategies. We document negative average returns for a broad range of strategy permutations over various parameter combinations and conclude that performance, if any, is due to randomness and not a verifiable mean reversion effect. Negligible tangency portfolio weights and high turnover further discourage from considering factor momentum based mean reversion as an active trading strategy in practice.

Keywords factor momentum · cross sectional momentum · time series momentum · mean reversion

1 Introduction

Quantitative investment strategies are systematic and rule-based in an effort to overcome behavioral biases and spontaneous, emotional decisions (*animal spirits*) that easily lead to digressions from the original strategy and can therefore result in sub-optimal investment behavior [Keynes, 1936]. These innately human errors in (financial) decision making led to the emergence of a new view on market efficiency that starkly contrasts Fama [1970]'s famous *Efficient Market Hypothesis* which states that "... prices always "fully reflect" available information ..." in an efficient market environment, making it *de facto* impossible and irrational trying to beat the market.

Shiller [1981], on the other hand, argues that stock price movements cannot be explained by new information about upcoming dividends (and earnings) only. Pedersen [2015] explains that Shiller's notion of *Inefficient Markets* leaves room for potential outperformance since price movements are thought to be not perfectly rational and can thus digress substantially from their intrinsic value through human errors and behavioral biases. The author finally brings some sort of consensus to both zealot perspectives and coins the term of *Efficiently Inefficient Markets*, arguing that markets are

*Our deepest gratitude goes to the whole staff of Quantumrock for welcoming us with open arms and always treating us as equals. Without their openness for the unusual format of a *project study* as well as their willingness to share the day to day work in a quantitative hedge fund with us, this research paper would not have been possible. Our special thanks goes to Michael Zeller, CIO of Quantumrock and our direct industry advisor, who did not only come up with the core idea behind our research topic and taught us valuable lessons about researching quantitative trading strategies, but also made the countless hours in the office so much more interesting and enjoyable. We would also like to thank our academic advisor, Tobias Kalsbach, for providing us with critical feedback throughout the project study, helping us understand academic financial research better in general and even taking time from his vacation to answer our questions. Lastly, we would like to express our thanks to Dr. Tarun Gupta (Invesco Ltd., ex AQR Capital Management LLC) who was readily available to help with technical inquiries regarding his own paper [Gupta and Kelly, 2018], one of our main references for this project study, as well as Dr. Matthias Hanauer (Robeco) for providing simple explanations to complex quantitative questions.

inefficient enough for professionals to achieve significant (out)performance in exchange for taking risk and digressing from the market portfolio while remaining *efficient* in a sense that the collective quest for *alpha* by all market participants pushes prices towards their fundamental values.

However, even in such an efficiently inefficient market environment, finding reliable signals in noisy financial time series data that can be detected, processed and acted upon by algorithms and mathematical models to generate robust returns continues to remain quite a challenge.

Factor investing is an established practice in quantitative asset management that aims at harvesting risk premia of well-documented and academically studied market anomalies. Put simply, a factor is constructed by utilizing a certain characteristic of a security, e.g. its market value, to sort an investment universe into groups of high/medium/low manifestations of said characteristic. Once an order has been determined, securities with the highest manifestations of this characteristic form the *winner* portfolio of the factor, whereas those securities with the lowest manifestations are part of the *loser* portfolio.[2] The factor's return is then calculated as the spread between the returns (equal-weighted) of these winner and loser portfolios. If there really is an effect based on this very characteristic, a long position in the winner portfolio and a short position in the loser portfolio should result in a statistically significant positive return.

Over the course of the financial research history, a plethora of such factors has been identified, so many in fact, that Cochrane [2011] even speaks of a *factor zoo* and that it has become its own discipline to propose models and frameworks for dealing with and selecting from the seemingly ever increasing number of (marginally significant and exotic) market anomalies.[3] One of the factors that consistently stands out positively in exhaustive robustness tests across time, markets, and asset classes is *individual stock momentum* [Chan et al., 2000, Asness et al., 2013, Ilmanen et al., 2021].

Jegadeesh [1990] was the first to document serial correlation in monthly security returns, suggesting the predictability of future returns via the security's own past returns and challenging the core assumption of the *Random Walk Theory* that changes in security prices are purely random and therefore impossible to forecast.[4] Whereas serial correlation coefficients at the 12-month level were positive and statistically significant (t-value of 9.09), those at the 1-month level were negative (t-value of -18.58), marking the birth of the short-term reversal (STR) factor. In a follow up study, Jegadeesh and Titman [1993] show that trading strategies going long in a portfolio of securities that have performed well over the past months (esp. for periods of six or 12 months) and going short in a portfolio of securities that have performed badly over the past months achieve statistically significant positive returns.

Carhart [1997] uses this momentum anomaly to extent Fama and French [1993]'s infamous three factor model to a four factor model, to explain the returns and evaluate the consistency in performance of mutual funds. Two years later, Moskowitz and Grinblatt [1999] expand the idea of momentum onto industries, arguing that a large portion of the abnormal returns generated by *individual stock* momentum can in fact be explained by *industry* momentum and that playing momentum on the industry level is a more profitable strategy than on the security level, at least for shorter investment horizons.

With this overwhelming academic evidence for its robustness, it is not surprising that momentum has become so popular in fundamental quantitative investing [Pedersen, 2015].

Recent research argues, however, that momentum in both, individual securities and industries, is in fact subsumed by different expressions of momentum on a more meta, *factor*, level. In their study over 15 prominent factors from 07/1963 to 12/2015, Ehsani and Linnainmaa [2019] conclude that momentum does not exist as a *standalone* factor *per se*, but rather "... is a manifestation of factor momentum", merely timing factors indirectly. Momentum in factors can be determined two-ways: on a cross-sectional basis, drawing a comparison over all factors in the sample and comparing their performance *relative* to each other and on a time-series basis, comparing individual factor returns to its own *absolute* past performance. Whereas this *cross-sectional* factor momentum (CSFM) is found to be able to explain abnormal returns of *industry* momentum, *time-series* factor momentum (TSFM) can explain returns of *individual stock* momentum [Ehsani and Linnainmaa, 2019, Arnott et al., 2021, Gupta and Kelly, 2018]. These studies find that TSFM outperforms CSFM on a risk-adjusted basis, attributing the superior performance to factor return autocorrelations.

[2]This is a very basic and hypothetical example. Usually, an assumption is made beforehand whether a high or low manifestation of the factor characteristic will result in positive expected returns. It is common practice, but not obligatory, to sort the manifestation of the characteristic that is expected to generate positive returns into the winner portfolio, and vice versa, in order to obtain a positive spread. Since Fama and French [1993], it has also become a standard to sort the investment universe by size first (big/small), and then after factor characteristic (high/medium/low), resulting in not two but six subportfolios.

[3]cf. e.g. Harvey et al. [2016], Feng et al. [2020].

[4]cf. e.g. Fama [1965].

Additionally, Gupta and Kelly [2018] show that factor momentum is not exclusive to the U.S. equity market, but can be detected all around the globe.

These exciting new findings and implications for active trading strategies did not go unnoticed by industry practitioners. For this project study, the CIO of the quantitative hedge fund Quantumrock hence asked us to explore whether an active stock-picking strategy building upon the idea of factor momentum and expanding it by a mean reversion approach could generate robust profitable returns. The general notion of *factor momentum* postulates that factors which have performed well recently (factor momentum winners) will continue to perform well in the future (i.e. their respective winner portfolio (Q5 quantile) will continue to perform well (long position) whereas their loser portfolio (Q1 quantile) will continue to perform badly (short position), leading to a positive Q5-Q1 spread), and factors that have performed badly recently (factor momentum losers) will also continue to perform badly in the future (i.e. their respective winner portfolio will continue to perform badly (long-position) whereas their loser portfolio will continue to perform well (short position), leading to a negative Q5-Q1 spread).

The main idea behind *factor momentum based mean reversion* now is that within those factors that recently performed well (winners), we expect securities in the top quantile (Q5) that performed the worst over the previous timestep to showcase an extraordinary performance in the next timestep, reverting towards the average return of the top quantile. For the best performing tickers of the past timestep in the bottom quantile (Q1), we expect a drop in performance in the next timestep, again reverting towards the average return of the quantile. The 'winner' portfolio of this mean reversion strategy is then made up of all the spreads between these Q5 'losers' (long position) and Q1 'winners' (short position) of each factor that has been determined a winner itself by either its cross-sectional momentum over all other factors in the sample or its very own time-series momentum. Analogously, we calculate the respective spreads for those factors that recently performed badly (losers), which make up the 'loser' portfolio of the strategy.

The final mean reversion strategy then trades the spread between the returns of the determined winner and loser portfolio, making it a self-financed, active trading strategy.

We test this hypothesis by first defining base and mean-variance optimized versions of cross-sectional and time-series factor momentum on a sample of 10 factors over the past 20 years (2001-2021) in the S&P 500 investment universe. Then, we form base cases of above mean reversion approach for each of the factor momentum variants (CSFM and TSFM), and perform a broad grid search on every tuneable parameter to find the best strategy permutations on a risk-adjusted basis.

Afterwards, we explore different tangency portfolio combinations of our various factor momentum and mean reversion strategies together with the Fama and French [2015] five factors as well as other popular anomalies. We consider the portfolio optimization process with and without short-selling restrictions, as well as with and without a proxy for a risk-free investment. Lastly, we perform some robustness tests and explore the impact of turnover and transaction costs on selected strategies.

Our results suggest that factor momentum based mean reversion does not exist as a robust and profitable anomaly within our underlying investment universe, factors and timeframe. The most basic versions of cross-sectional and time-series factor momentum based mean reversion (MRCSFM and MRTSFM) earn negative Sharpe ratios of -0.21 and -0.34. We perform a broad grid search over various parameters but document negative average returns for every strategy permutation. Even those versions with the 'optimal' underlying parameter values, arguably overfitted to the timeframe, earn Sharpe ratios of -0.12 and -0.07. Linear regressions over the S&P 500 and different factor models result in negative alphas and tangency portfolio weights are either zero or negative to finance long positions in other portfolio constituents.

Since average returns of the mean reversion strategies are statistically unsignificant, simply shorting them is not an advisable strategy to generate reliable profits and any performance has to be attributed rather to randomness than any existing effect. Our robustness checks reveal high volatilities in returns and large drawdowns, as well as underperformance to the market in terms of excess returns. Furthermore, the mean reversion strategies require a lot of trading, quickly accumulating transaction costs.

Taken together, betting on reversion to the mean within factor momentum does not seem to be a profitable trading strategy. Nevertheless, our results are prone to many limitations and we encourage further research with larger datasets and more sophisticated signal generation methods.

2 Data

2.1 Asset Universe

Our asset universe comprises S&P 500 constituents over the time period of 2001-06-11 to 2021-06-11.[5] We account for the *survivorship bias* by only considering those tickers that have actually been S&P 500 constituents for a specific timestep. Thus, at any time, we work with around 500 tickers and not with the complete historical S&P 500 constituents universe of around 924 tickers.

2.2 Factor Sample

We use a total of 10 factors that have been constructed from *FactSet* data, provided to us by Quantumrock's Heads of AI Systems and Algorithmic Trading. Table 2.1 lists said factors and groups them into the commonly known risk factors of *Size*, *Value*, *Quality* and *Dividends*.

Table 2.1: Factor Sample Description

Category	FactSet Acronym	Description
Size	FREF_MARKET_VALUE_COMPANY	Market Value
Value	FF_PE	Price to Earnings
	FF_PSALES	Price to Sales
	FF_PBK	Price to Book
Quality	FF_ROA	Return on Average Assets
	FF_ROTC	Return on Average Invested Capital
	FF_ROE	Return on Equity
Dividends	FF_DIV_YLD	Dividend Yield
	FF_DPS	Dividends per Share
Other	FSI_DAYS_ANY_EXCHG	Short Interest Ratio

2.3 Factor Formation & Portfolio Returns

Each ticker in our underlying data set features several factor characteristics at every time step, from which we form factors. We encounter some *NA* values since not all tickers are traded on exactly the same days. These missing values are filled with those of the preceding day, except for the indicator field that gives binary information on whether that ticker is part of the S&P 500 index on that particular trading day (1) or not (0). In these cases, we fill the missing indicator value with 0 to avoid the risk of including a ticker into our calculations that has not been part of the S&P 500 on that particular day.

It has become a common practice in financial academic research to split the data set first by size into bins with high and low market capitalization, respectively, before splitting each size bin further by factor characteristic (low/medium/high) [Fama and French, 1993]. Since we are working on a relatively small dataset of predominantly large and liquid companies[6] and not on the whole universe of U.S. equities, we refrain from forming portfolios in this *2x3* manner. Instead, we build factor portfolios by sorting the tickers after the respective return predictor (e.g. market value for size, ROE for quality) and then determining the top and bottom *quantiles* (Q5 and Q1 respectively) on a monthly basis. Factor returns are then calculated as the spread between the returns (equal-weighted) of the Q5 (long) and Q1 (short) quantiles.[7] Figure A.1 presents the construction of the respective factor portfolios graphically. We also do not lag accounting information [Fama and French, 1992, 1993], since the per-day accounting information for each ticker was already included in the original FactSet data. As it is common practice, we rebalance these portfolios monthly.

Factor returns are calculated ex-post per timestep, starting 07/2001, and represent the *simple* return this factor would have theoretically achieved over the respective timeframe. Note that this factor return could not be realized in practice since we would not know ex-ante which tickers would ultimately make up the factor (would be part of the Q5 and

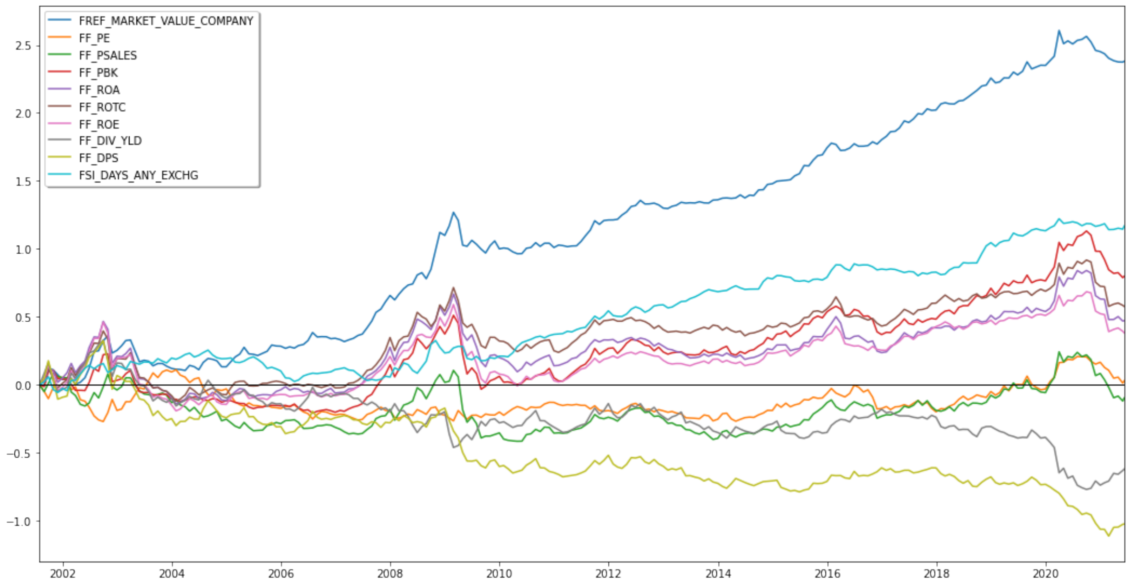
[5] YYYY-MM-DD format.

[6] cf. `https://www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-us-indices.pdf`.

[7] We actually did form factor portfolios both ways but did not find any real performance difference when using one over the other, hence sticking with the more straightforward solution.

Q1 quantiles) at the end of the timeframe. We thus only use these artifical returns for signal generation of the factor momentum and later mean reversion strategies. Figure 2.1 displays the cumulative log returns of our 10 factors over the whole time period.

Figure 2.1: Cumulative Factor Log Returns



2.4 Factor Statistics

2.4.1 Return Distributions

Going forward, we will be using monthly log returns since they closely resemble a normal distribution with the exception of some fat tails (see Figure A.2), which, however, are commonly found in financial time series data [Chavez-Demoulin et al., 2011]. In contrast to *simple* returns, *log* returns are time-additive, which facilitates the calculation of cumulative returns, as well as annualized expected return metrics [Dorfleitner, 2003].

2.4.2 Summary Statistics

Figure 2.2 explores the correlation between the factor returns. As indicated by color ranging from dark red (inversely correlated) over light yellow (no correlation at all) to dark green (perfectly correlated), we can determine that some of the factors are closely correlated with each other such as FREF_MARKET_VALUE_COMPANY, FF_PSALES, FF_PBK, FF_ROA, FF_ROTC and FF_ROE, which makes sense, given that these are mainly calculated from pricing data. Other factors such as FF_PE, FF_DIV_YLD and FSI_DAYS_ANY_EXCHG seem to be negatively or completely uncorrelated to the rest, which motivates their combination with respect to strategy/portfolio construction.

We report expected annual returns, standard deviations, Sharpe and Sortino ratios [Sharpe, 1966, 1975, 1994, Sortino and van der Meer, 1991, Sortino and Price, 1994] as well as Fama and French [2015] five factor alphas[8] for every factor in Table 2.2. Figure 2.3 additionally shows corresponding error terms and t-values for these metrics.[9] We decided to include the Sortino ratio since it zooms in on the *downside* risk of returns, while the Sharpe ratio does not differentiate between *positive* and *negative* return fluctuations. Also, we assume the risk-free rate to be zero which is not too far-fetched, given that the one-month U.S. Treasury Bill rate has been 0.10% on average over the past 20 years.[10]

[8]Fama/French Factors are fetched from the Kenneth R. French website (`https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`).

[9]We determine confidence intervals for Sharpe and Sortino ratios after the methodology described by Lo [2003].

[10]Fama and French [1993] and the Kenneth R. French website state that the one-month U.S. Treasury Bill rate is taken as a proxy for the risk-free rate.

Figure 2.2: Factor Return Correlations

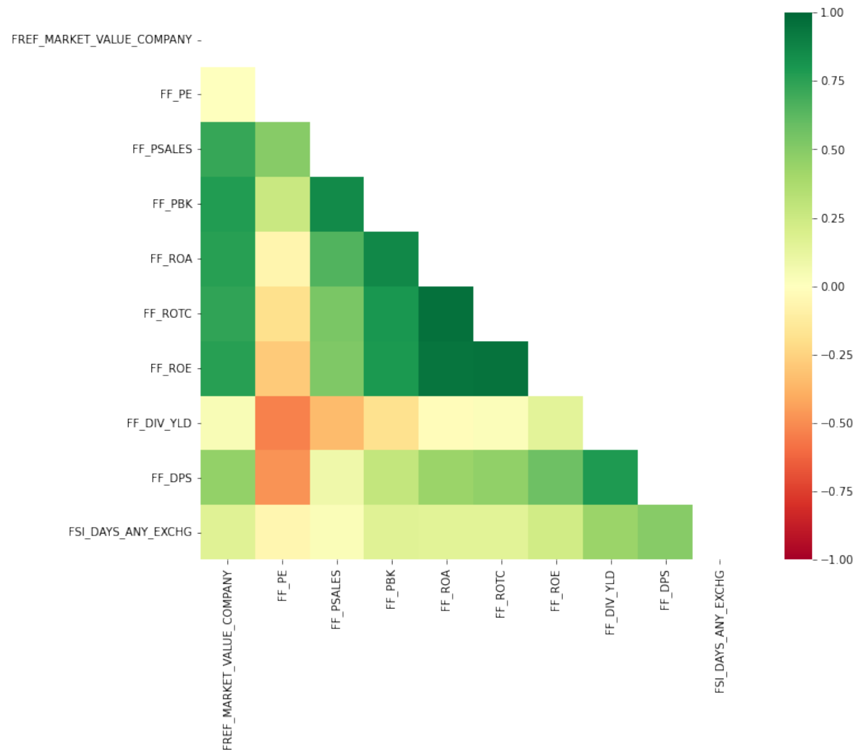
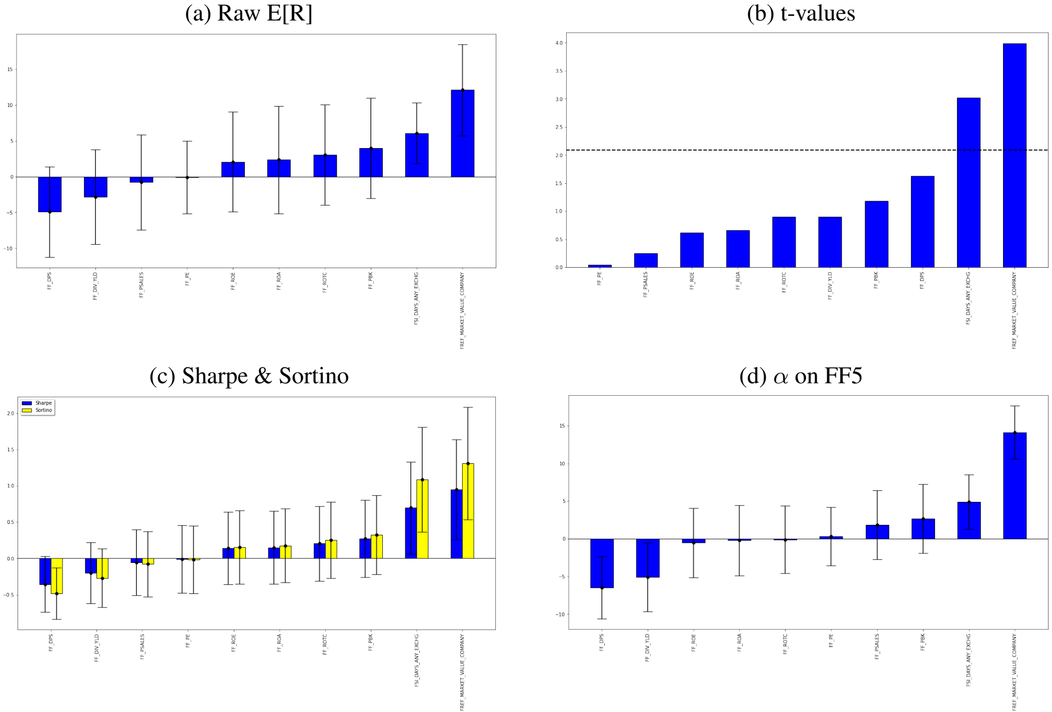


Table 2.2: Factor Sample Statistics

Factor	E[R]	STD	Sharpe	Sortino	FF5 α
FREF_MARKET_VALUE_COMPANY	12.08*	13.36	0.90*	1.23*	14.12*
FF_PE	-0.12	10.83	-0.01	-0.02	0.32
FF_PSALES	-0.80	14.25	-0.06	-0.07	1.82
FF_PBK	3.96	15.15	0.26	0.30	2.66
FF_ROA	2.36	16.33	0.14	0.16	-0.22
FF_ROTC	3.01	15.16	0.20	0.23	-0.11
FF_ROE	2.06	15.35	0.13	0.14	-0.53
FF_DIV_YLD	-2.85	14.33	-0.20	-0.25	-5.08*
FF_DPS	-4.92	13.72	-0.36	-0.45*	-6.49*
FSI_DAYS_ANY_EXCHG	6.07*	8.92	0.68*	1.05*	4.87*

Rounded to two decimals. Superscript of * denotes statistical significance at p = 0.05 level. E[R], STD and FF5 α are reported in percent.

Figure 2.3: Factor Performance



(a) Raw E[R] (b) t-values (c) Sharpe & Sortino (d) α on FF5

Six out of the 10 factors exhibit positive average expected returns, with only two of those being statistically significant at 95% confidence. Sharpe and Sortino ratios are predominantly positive aswell, with 0.90 and 1.23, respectively, being the highest values. This is already a quite respectable result, given the uncomplicated construction of the factor portfolios. FF5 αs are statistically indistinguishable from zero for most factors, but for those with p-values below 0.05, we find expressive positive/negative numbers.

Note, that we orient the long/short legs in the same manner for each factor, sorting the underlying factor characteristic from large to small (e.g. for FREF_MARKET_VALUE_COMPANY, we sort tickers in descending order after their market capitalization, resulting in the largest companies to fall into the Q5 quantile and the smallest into the Q1 quantile).[11] Here, the traditional notion of small companies outperforming large companies[12] does not seem to hold. This observation is in accordance with recent research. In Fama and French [2015], the authors acknowledge that, amongst others, the size factor struggles to capture the returns of large growth companies - characteristics that we can typically observe in S&P 500 companies. In a very recent study by Blitz and Hanauer [2020] about the present relevance of the size factor, the authors conclude that its premiums are statistically unsignificant when observed by itself, making it a weak factor on a stand-alone basis.

2.4.3 Autocorrelations

Following the example of Avramov et al. [2017], Gupta and Kelly [2018], Ehsani and Linnainmaa [2019] who examine cross-sectional and time-series factor momentum, we investigate potential serial correlation in the returns of our factor sample. To make sure that there is no linear trend underlying our data, we first test the returns for stationarity. Stock returns are generally distributed around zero which already points to stationarity, nevertheless, we test our factor returns with the Augmented Dickey-Fuller test (ADF)[13] which returns large and negative test statistics with very small and hence significant p-values.[14] We can therefore safely assume stationarity in our underlying factor returns and continue

[11] At first glance, this signing convention might go against the traditional notion of assigning the long leg to the quantile that is expected to have positive returns and the short leg to the opposite quantile with negative expected return for each factor separately. However, since the subsequent strategies that build upon these factor returns are all momentum strategies that exclusively get their trading signals from past factor returns, the signing notation does not matter. See also Arnott et al. [2021].

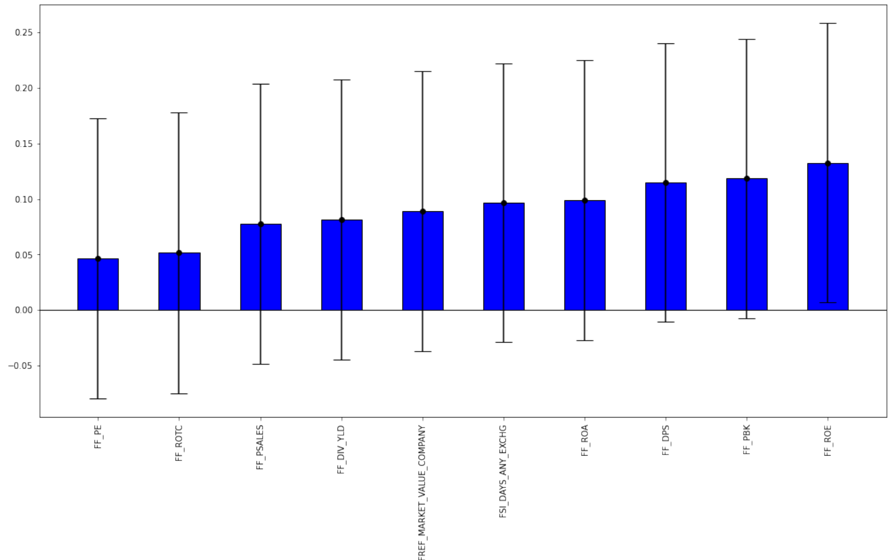
[12] cf. e.g. Banz [1981], Fama and French [1992, 1993].

[13] cf. Dickey and Fuller [1979].

[14] Please refer to Table A.1 for the exact figures.

with examining monthly first-order autoregressive coefficients (AR(1)), as displayed in Figure 2.4.

Figure 2.4: Factor Sample Monthly AR(1) Coefficients

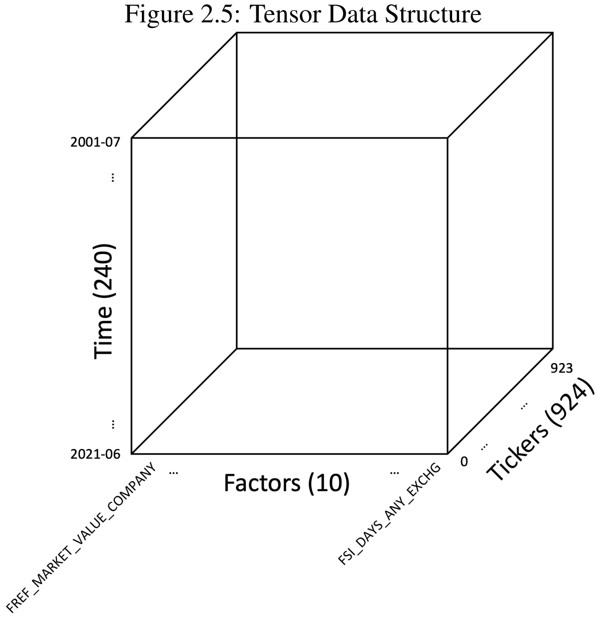


All AR(1) coefficients are positive with the average being 0.09. However, it is only for the FF_ROE factor, that we can statistically confirm its AR(1) coefficient at a 95% confidence level. Unfortunately, unlike above authors, we cannot report robust factor serial correlation in our sample of factors. These results make it seem improbable, but not entirely impossible, to time factors via their own past returns.

2.5 Tensor

We store the post-processed data in a 3D tensor for easy access through Python's *numpy* and *pandas* packages with the following dimensions: *tickers* (924) x *time* (240 months) x *fundamental & supplementary data* (16), which consists of 10 factors and six information points. The resulting data structure can be imagined as illustrated in Figure 2.5.

Figure 2.5: Tensor Data Structure



3 Benchmarks & Factor Momentum

Before we turn to the actual mean reversion strategies, we first have to define the underlying factor momentum approaches as well as some benchmarks to compare them to. Returns of these strategies start from 08/2001 since they rely on previous month factor returns. We start by introducing the concepts of cross-sectional and time series factor momentum before moving on to discussing their statistics.

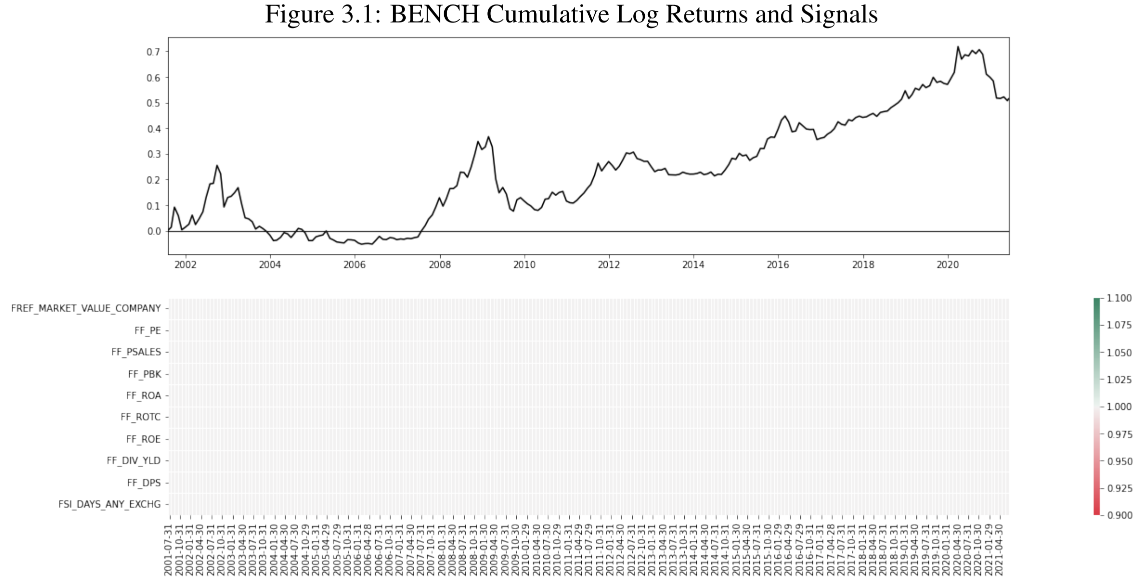
3.1 Market Portfolio (SP500)

Since our ticker universe comprises solely S&P 500 constituents, taking the S&P 500 index itself as a proxy for the market portfolio seems a reasonable choice. We fetch historical data for ˆGSPC from Yahoo Finance.[15]

3.2 A naïve benchmark (BENCH)

In order to determine whether the effort of cherry-picking factors based on their cross-sectional or time-series momentum pays off, we follow Avramov et al. [2017], Ehsani and Linnainmaa [2019]'s example and define a naïve benchmark strategy that goes long in every factor at the beginning of the time period and holds its positions (equal weights, 1/N rule) until the end. Note that going *long* in a factor still encompasses a long leg (Q5) and a short leg (Q1). Just like described in section 2.3, these benchmark returns are yet again only theoretical in nature and not achievable in practice, since we do not know at the beginning of each time step which tickers would ultimatively end up in the top and bottom quantiles of the respective factor at the end of the time step. Figure 3.1 displays the theoretical cumulative log returns of the BENCH strategy as well as its monthly trading signals (grey (long)) in every factor.

Figure 3.1: BENCH Cumulative Log Returns and Signals



3.3 Cross Section Factor Momentum (CSFM)

We model a simple version of CSFM after the example of Ehsani and Linnainmaa [2019], Arnott et al. [2021].[16] As the most basic, *plain vanilla* (PV), version, CSFM_PV features a formation period of one month and a cutoff value of 0.5. Thus, for each month t, we first cross-sectionally determine the median return value across all factors. The winner portfolio in $t+1$ is then made up of those factors whose returns in t lied above the median, and the loser portfolio respectively holds the factors whose returns underperformed the median. Both portfolios are formed on an equal-weight basis. CSFM_PV then goes long in the winners and short in the losers, effectively trading the spread between the

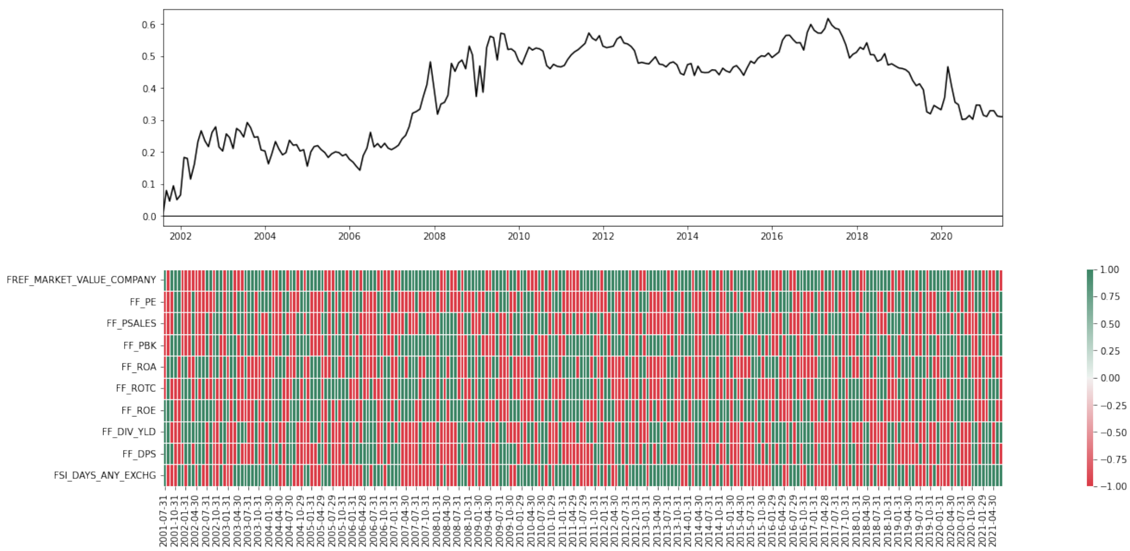
[15] `https://finance.yahoo.com/quote/%5EGSPC?p=^GSPC&.tsrc=fin-srch`.

[16] Unlike the theoretical factor returns from above, the returns of this factor momentum strategy can actually be realized in practice. The strategy is structured in a way that it excludes the indirect *look-ahead bias* (we do not know ex-ante which securities will end up in the Q5 and Q1 quantiles of each factor at month's end) underlying the factor returns.

returns of these subportfolios for month *t+1*.[17] It is worth noting that for CSFM, winner and loser portfolios always hold the same amount of factors, except if one particular factor return exactly equals the median value. In this case, said factor will not be part in either portfolio and the weights are adapted accordingly such that they still sum up to one for each portfolio. CSFM strategies are formed on a unit leverage basis, i.e. the short positions finance the long positions without the need of an additional investment. All portfolios are rebalanced monthly (holding period of one month).

Figure 3.2 displays cumulative log returns as well as monthly trading signals (dark red (short), grey (no position), dark green (long)) across all factors for the CSFM_PV strategy.

Figure 3.2: CSFM_PV Cumulative Log Returns and Signals



We can clearly see that CSFM_PV allocates factor weightings equally across all factors, always going long in one half of the sample and shorting the other. Positions in a particular factor are generally held over various months, confirming that the strategy actually tries to play factor momentum as we defined it above. CSFM_PV has an expected annual return of 1.55% (t-value of 0.61) with a standard deviation of 11.41% (Sharpe ratio of 0.14). We will go into more detail about CSFM_PV's statistics and the performance of the individual long and short legs shortly.

The works that have guided us so far in the conception of this cross-sectional factor momentum strategy perform sensitivity analysis on formation and holding periods but leave the cutoff value that determines whether a factor is a cross-sectional over- or underperformer at the median only [Ehsani and Linnainmaa, 2019, Arnott et al., 2021]. For this reason, we decided to choose the cutoff value as a flexible parameter for further analysis, fixating the holding period at one month to keep the notion of monthly rebalancing. Doing so allowed us to examine whether more extreme selections of cross-sectional winners and losers, e.g. the 20% best and worst factors with respect to the cross section of all considered factors, achieve higher expected returns.

Figure 3.3 summarize the results of our grid search over various formation periods (1, 3, 6, 9 and 12 months)[18] as well as cutoff values (0.1, 0.2, 0.3, 0.4 and 0.5) while keeping the holding period constant at one month.

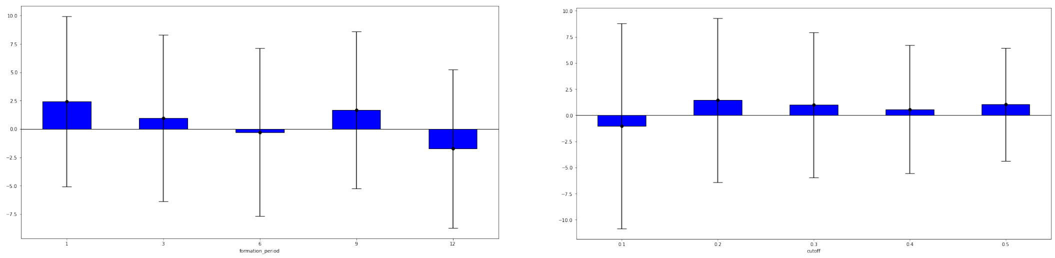
The average returns do not seem to follow a clear pattern. We observe higher expected returns for very small formation periods of one month, but also for moderately long ones of nine months. For six and 12 months, we obtain negative results. Cutoff values above 0.1 yield positive expected returns, i.e. extreme percentile selections (90% and 10% percentile, cutoff = 0.1) do not affect returns positively. Note that all 95% confidence intervals are centered around zero, pointing towards non-robustness of the reported results. Panel (a) in Table A.2 reports the exact figures and t-values for

[17] See Panel (a) in Figure A.3 for a graphical representation of the returns of CSFM_PV and its respective winner and loser portfolios.

[18] For formation periods longer than one month, we calculate average returns and compare those against the cutoff.

Figure 3.3: Average Annual E[R] for various Formation Periods and Cutoff Values (CSFM)

(a) Formation Periods (b) Cutoff Values

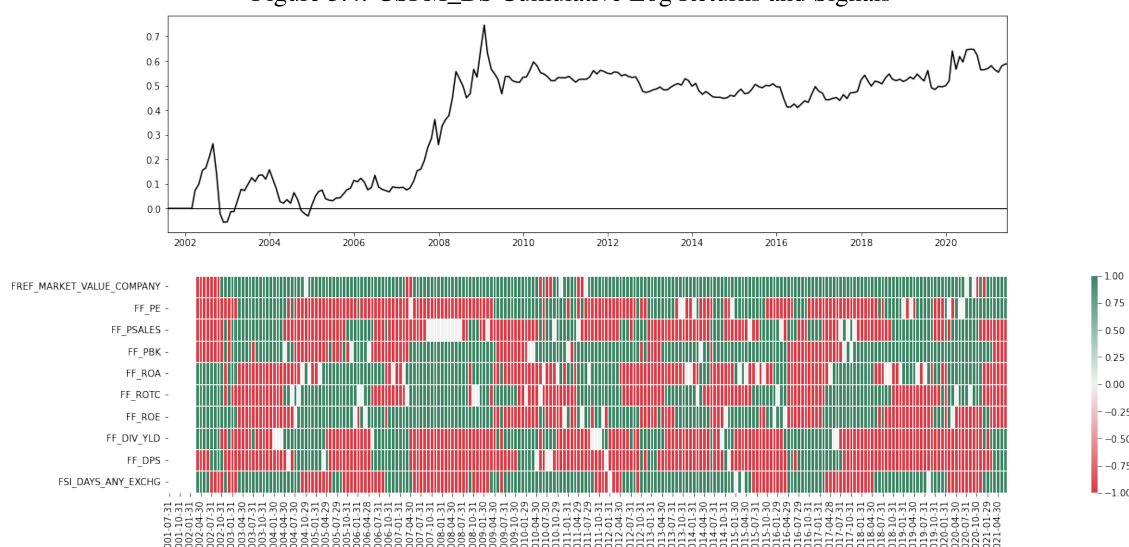


this sensitivity analysis and Panel (a) in Figure A.4 shows that the observed effects also hold true for average Sharpe ratios.

Going forward, we will consider the CSFM strategy version with the highest overall risk-return profile (i.e. *best Sharpe*, BS) as an additional benchmark (CSFM_BS). In our case, sorting factors after their nine months past average returns (formation period of nine months) and going long/short in those with above/below median returns (cutoff of 0.5) yielded an expected annual return of 2.94% (t-value of 1.13) and a standard deviation of 11.63%, leading to a Sharpe ratio of 0.25.[19] On a sidenode, CSFM_PV actually turned out to be the minimum variance version of the grid search results.

Figure 3.4 displays yet again cumulative log returns as well as monthly trading signals across all factors for the CSFM_BS strategy.

Figure 3.4: CSFM_BS Cumulative Log Returns and Signals



For CSFM_BS, positions in the individual factors are held for much longer time periods than for CSFM_PV taking more advantage of up- and downtrends. At times, this strategy invests in only nine out of our 10 factors, which can happen if the average return of one factor happens to be the cross-sectional median value, in which case no position is taken in this very factor.

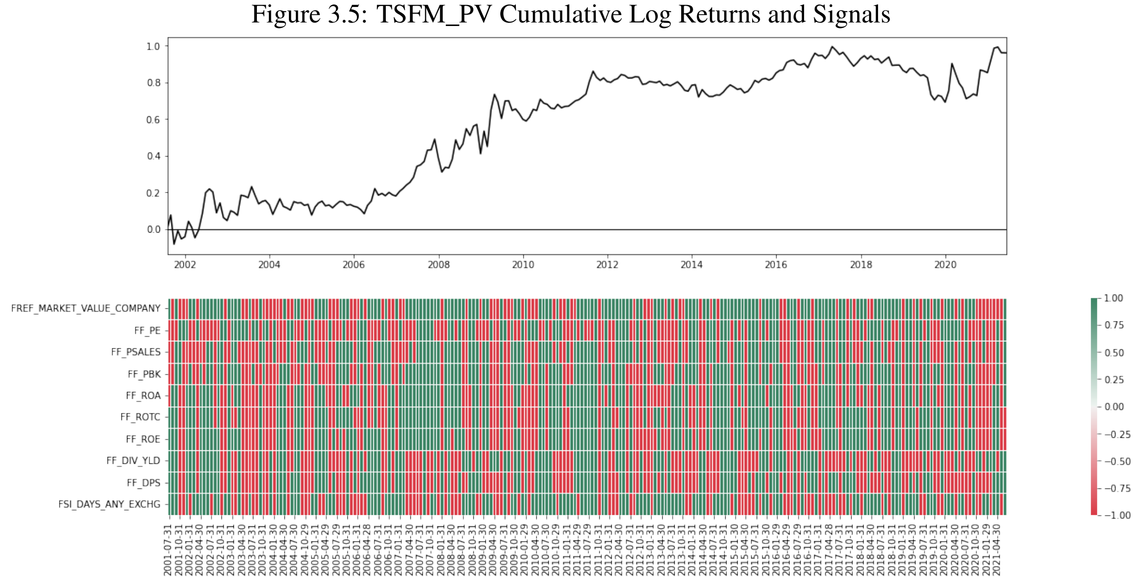
[19]See Panel (b) in Figure A.3 for a graphical representation of the returns of CSFM_BS and its respective winner and loser portfolios.

3.4 Time Series Factor Momentum (TSFM)

Analogous to CSFM, we took inspiration for modeling a simple version of TSFM from Ehsani and Linnainmaa [2019], Gupta and Kelly [2018].[20] We yet again define a *plain vanilla* version, TSFM_PV, characterized by a formation period of one month and a cutoff value of zero. Note that we are now not looking at the cross-section of factor returns anymore but at each factor's own time series of returns, in order to determine whether to go long or short. Each month t, we determine for each factor separately, whether its returns in t were positive (above the cutoff of zero), or negative (below the cutoff of zero). The winner portfolio in $t+1$ is then made up of those factors with positive returns in t, and the loser portfolio respectively holds the factors with negative returns. Both portfolios are formed on an equal-weight basis. As with CSFM_PV, TSFM_PV then buys the winners and sells the losers, trading the spread between their returns for month $t+1$.[21] Unlike for CSFM, winner and loser portfolios for TSFM do not necessarily have to hold the same amount of factors. TSFM_PV will go long in every factor with past returns above zero, which, in bullish market phases, can include every factor in the sample. Accordingly, in extreme market downturns, TSFM_PV might go short in every factor, solely depending on whether its past return was negative or not. To compensate for these fluctuations, we make sure that winner and loser portfolios distribute weights equally and always sum up to one, such that TSFM_PV is yet again formed on unit leverage. All portfolios are rebalanced monthly (holding period of one month).

Figure 3.5 displays cumulative log returns as well as monthly trading signals across all factors for the TSFM_PV strategy.

Figure 3.5: TSFM_PV Cumulative Log Returns and Signals



In accordance with above definition, TSFM_PV does not strictly allocate factor weightings equally across all factors. Especially in times leading up to market crashes such as the pre 2008 era for instance, TSFM_PV disproportionally goes long in the majority of factors, confirming that the strategy works as intended. On average, TSFM_PV goes long in 5.5 factors, showing a slight long-bias. Over the whole time period, we observe an expected annual return of 4.80% (t-value of 1.42) with a standard deviation of 15.12% (Sharpe ratio of 0.32) for TSFM_PV. More detailed statistics will be discussed shortly.

As it was the case with the CSFM strategy, previous work that we build upon perform sensitivity analysis on formation and holding periods only, yet again leaving the cutoff value untouched at zero [Ehsani and Linnainmaa, 2019, Gupta and Kelly, 2018]. We aim to fill this gap and therefore choose the cutoff value yet again as flexible variable for further

[20] Gupta and Kelly [2018] actually define a more sophisticated version of TSFM that scales factor exposures by the inverse of the annualized three to 10 year factor volatility. While certainly a more robust approach, we decided to lean our implementation of TSFM less on this sophisticated version and rather on the simple version proposed by Ehsani and Linnainmaa [2019] since our timeframe of 20 years already is comparably short. Taking the minimum of three years factor volatility into consideration would further shorten our timeframe which we thought not to be a sustainable idea.

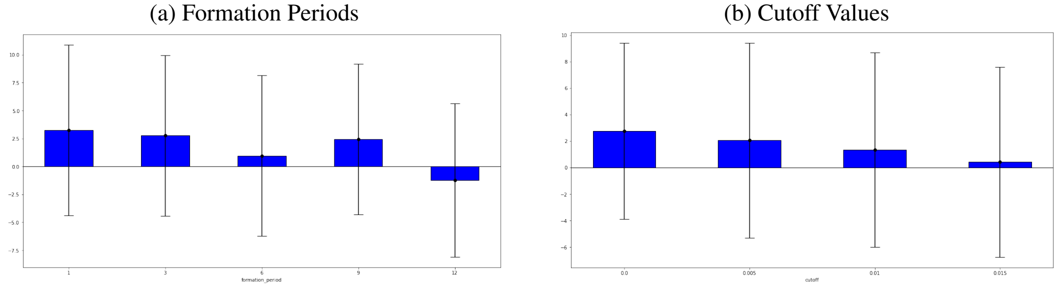
[21] See Panel (c) in Figure A.3 for a graphical representation of the returns of TSFM_PV and its respective winner and loser portfolios.

analysis while keeping the holding period constant at one month. It might be interesting to see whether more extreme selections of time-series winners and loser, e.g. factors with past returns above/below a threshold of 1% (instead of 0%), achieve higher expected returns.

The results of our grid search over various formation periods (1, 3, 6, 9 and 12 months)[22] as well as cutoff values (0, 0.005, 0.01 and 0.015)[23] is summarized in Figure 3.6. Holding periods are kept constant at one month.

Figure 3.6: Average Annual E[R] for various Formation Periods and Cutoff Values (TSFM)

(a) Formation Periods (b) Cutoff Values



Again, we cannot find clear patterns in average returns. We observe higher expected returns for very small formation periods of one and three months, but also for moderately long ones of nine months. Returns for six months are close to zero and for 12 months, we obtain negative results. For the cutoff values, there seems to be the notion that higher cutoff values, i.e. more extreme selections of winners and losers, do lead to smaller expected returns. Unfortunately, all of our calculations are yet again non-significant at 95% confidence, so we can neither confirm nor rule out possible effects. Panel (b) in Table A.2 reports the exact figures and t-values for this sensitivity analysis and Panel (b) Figure A.4 shows that the observed effects also hold true for average Sharpe ratios.

Analogous to CSFM_BS, we define a *best Sharpe* version, TSFM_BS, as an additional benchmark. Coincidentially, TSFM_PV turned out to be optimal in terms of the Sharpe ratio already, which is why TSFM_PV and TSFM_BS actually describe the same strategy that goes long/short in factors whose past month's return was above/below zero.[24] We therefore refrain from printing the factor portfolio decomposition and signal figures twice.

3.5 Factor Momentum Statistics

Now that we have defined all plain vanilla and best Sharpe versions of the factor momentum strategies, this section, analogously to section 2.4.2, reports expected annual returns, standard deviations, Sharpe and Sortino ratios for each of them, alongside their respective winner and loser portfolios in Table 3.1 and Figure 3.7.

The naïve benchmark BENCH that follows a simple, equal-weighted buy and hold strategy of every factor over the whole timeframe earns on average 2.57% a year (t-value of 1.26) with a standard deviation of 9.16%. This result demonstrates that the risk of investing in individual factors (average volatility of 13.74%) is mitigated by an equal-weight combination approach.

CSFM_PV actually does not manage to outperform this simple approach, earning only around 60% of BENCH's returns at a 26% higher standard deviation, resulting in a Sharpe ratio effectively half of the benchmark's. Although TSFM_PV features average annual returns of 4.80% and thereby slightly outperforms BENCH, this comes at the cost of higher volatility (15.12% standard deviation), resulting in a Sharpe ratio very close to that of the benchmark. Drawing on the Sortino ratio, however, reveals that not all of this extra volatility is necessarily negative - with a figure 34% higher than that of the benchmark, the time-series factor momentum approach introduces more upside volatility, detrimental for the Sharpe ratio, but welcome in reality. The best Sharpe versions of CSFM and TSFM perform similarly to their plain vanilla counterparts, barely outperforming BENCH at the cost of more (upside) volatility.

[22] As for CSFM, we calculate average returns in the case of formation periods longer than one month, and compare those against the cutoff.

[23] Standing for 0%, 0.5%, 1% and 1.5%, respectively.

[24] To stay consistent with the two CSFM benchmarks, we will still consider TSFM_PV and TSFM_BS separately and not merge them together. However, please stay alert that both strategies are actually the same.

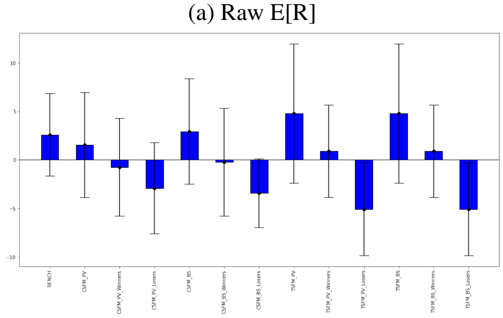
Table 3.1: Factor Momentum Statistics

Factor	E[R]	STD	Sharpe	Sortino
BENCH	2.57	9.16	0.28	0.33
CSFM_PV	1.55	11.41	0.14	0.21
Winners	-0.76	10.76	-0.07	-0.10
Losers	-2.91	10.40	-0.28	-0.27
CSFM_BS	2.94	11.63	0.25	0.32
Winners	-0.24	12.17	-0.02	-0.02
Losers	-3.42	7.69	-0.44*	-0.50*
TSFM_PV	4.80	15.12	0.32	0.47
Winners	0.91	10.16	0.09	0.13
Losers	-5.09*	10.58	-0.48*	-0.49*
TSFM_BS	4.80	15.12	0.32	0.47
Winners	0.91	10.16	0.09	0.13
Losers	-5.09*	10.58	-0.48*	-0.49*

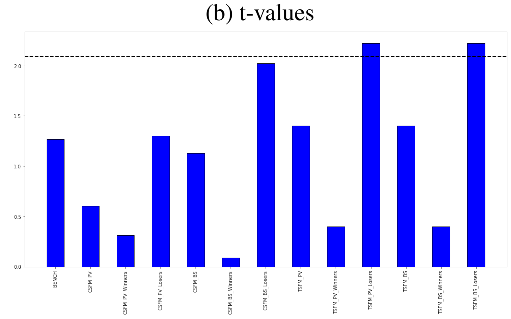
Rounded to two decimals. Superscript of * denotes statistical significance at p = 0.05 level. E[R] and STD are reported in percent.

Figure 3.7: Factor Momentum Statistics

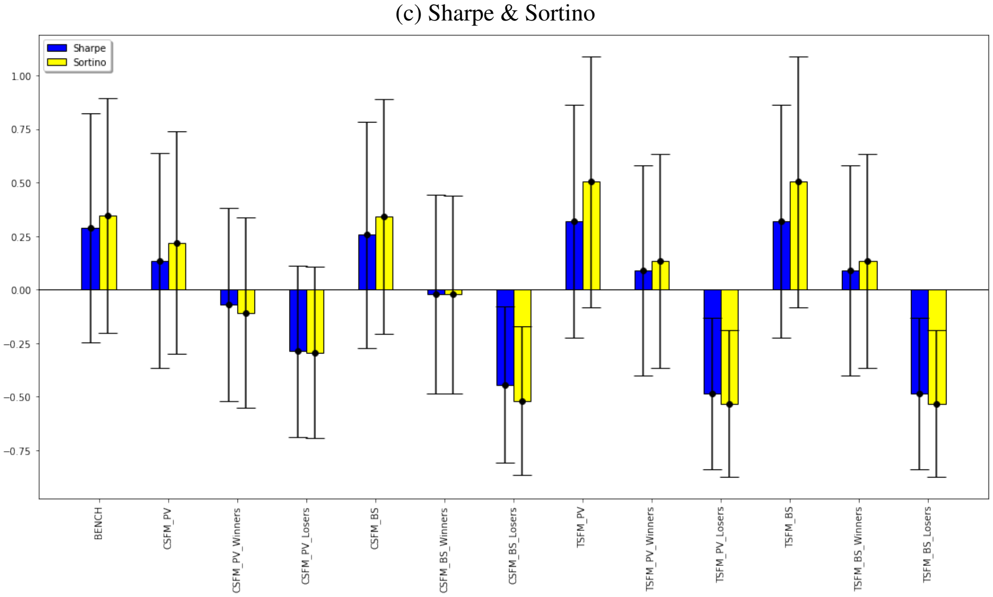
(a) Raw E[R]



(b) t-values



(c) Sharpe & Sortino



Drilling further into the respective winner and loser portfolios reveals that most of the factor momentum returns actually stem from their loser portfolios. Since the strategies go short in the loser portfolio, their negative returns actually contribute positively to the overall strategy performance. These are also the only cases where we can partially statistically confirm our results on 95% confidence. This observation stands in stark contrast to the results of Ehsani and Linnainmaa [2019] who find small returns statistically indistinguishable from zero for their loser portfolios while reporting very significant (t-values around 9) positive expected returns for their winner portfolios.

Loser portfolios of our factor momentum strategies go short in the worst performing factors within the sample, regardless of whether they have been determined as such via a cross-sectional or a time-series approach. These results might suggest that the most profitable strategies aim at premia which, in reality, are locked behind short-sale restrictions [Avramov et al., 2017]. In contrast to Avramov et al. [2017], however, who also report larger and significant returns for their short-leg portfolios, our loser portfolios are still made up of a long leg and a short leg (the shorted factors consist of long and short legs themselves) and do not only trade short-leg losers. Hence, we cannot attribute the performance of the loser portfolio to unattainable shorting premia only.

A possible explanation for these mostly unsignificant results may lie in the fact that neither factor returns nor autocorrelations, a crucial prerequisite for factor momentum, are statistically significant for the majority of our factors.

3.6 Benchmark Statistics

To conclude this section, we want to take a brief look into the statistics of our final sample of benchmarks. Table 3.2 explores the correlation between SP500 as a market proxy, BENCH as a buy-and-hold strategy, as well as plain vanilla and best Sharpe versions of cross-sectional and time-series factor momentum.

Table 3.2: Benchmark Return Correlations

	SP500	BENCH	CSFM_PV	CSFM_BS	TSFM_PV	TSFM_BS
SP500	1.00					
BENCH	-0.66	1.00				
CSFM_PV	-0.04	0.01	1.00			
CSFM_BS	-0.00	-0.05	0.23	1.00		
TSFM_PV	0.02	-0.11	0.86	0.23	1.00	
TSFM_BS	0.02	-0.11	0.86	0.23	1.00	1.00

The naïve benchmark is moderately negatively correlated to the S&P 500, whereas the factor momentum strategies are not correlated at all to the market or the naïve benchmark (values very close to and around zero). Amongst themselves, the plain vanilla versions of CSFM and TSFM are closely correlated (0.86), an effect that, although even more extreme, has already been discovered by Gupta and Kelly [2018]. The best Sharpe versions of CSFM and TSFM, however, are only slightly correlated to each other (0.23).

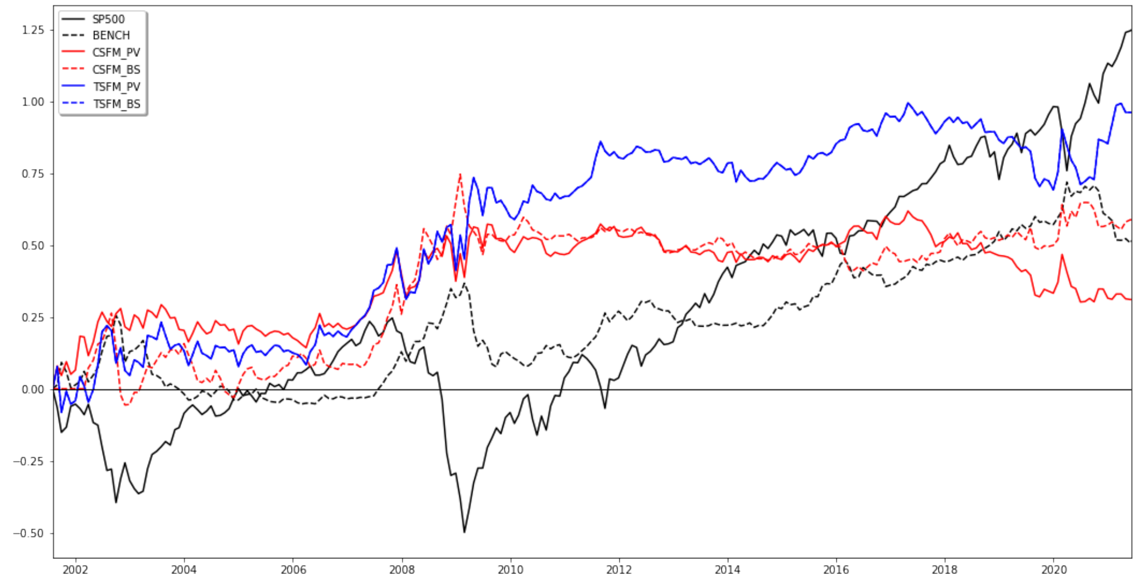
Figure 3.8 visualizes the cumulative log returns of all benchmarks over the whole time period.

With an average annual return of 6.24% (t-value of 1.87) and a standard deviation of 14.94% (Sharpe ratio of 0.42), the S&P 500 performs the best over the observation frame of 20 years amongst all benchmark strategies. These returns, however, are tarnished by large drawdowns in the post 2002 era and during the financial crisis of 2008/09. It is especially in these times of extreme market turmoil that the factor momentum based strategies shine. Their overall negative or non-relatedness to the market aids them in not only overcoming these bear markets, but also, at least in the case of CSFM and TSFM, in achieving respectable positive returns instead. Interestingly enough, at the maximum drawdown of the S&P 500 in early 2009, CSFM and TSFM which have been moving in a parallel fashion until then, start separating, with the CSFM versions moving rather sideways over the rest of the timeframe, completely missing out on the grand comeback of the market, whereas TSFM manages to capture at least some of the advance and ends up trading significantly higher than its cross-sectional counterparts at the end of the time period.

We feel the need of addressing the lack of statistically significant results for the presented strategies. The academic literature around the topic of factor momentum[25] reports large and impressive t-values for their various factor momentum

[25] Mainly Avramov et al. [2017], Gupta and Kelly [2018], Ehsani and Linnainmaa [2019], Arnott et al. [2021].

Figure 3.8: Benchmark Cumulative Log Returns



strategies. We have to stress, however, that our sample merely consists of 10 factors with around 240 data points each (monthly returns over 20 years), whereas the samples of our academic predecessors span up to 65 factors with starting dates as early as 1963. While surely delivering a more realistic image of the modern stock market, the post 2000 era is generally known for diminishing anomaly returns with structural and technological advances being the main drivers [Chordia et al., 2011, 2014, Avramov et al., 2017].[26] We do not question the results of aforementioned authors, but we argue that the pre 2000 era possibly contributed significantly to the stellar performance and robustness of their factor momentum strategies.

4 Factor Momentum based Mean Reversion

With the previous sections having established a sound baseline and the necessary foundations, this section finally goes into the details of the mean reversion strategies based on factor momentum. Since we differentiate between two versions of factor momentum (CSFM and TSFM) and build our strategies on top of both, mean reversion based on cross-sectional factor momentum will be abbreviated as MRCSFM, whereas mean reversion based on time-series factor momentum will be encoded as MRTSFM going forward.

4.1 MRCSFM and MRTSFM

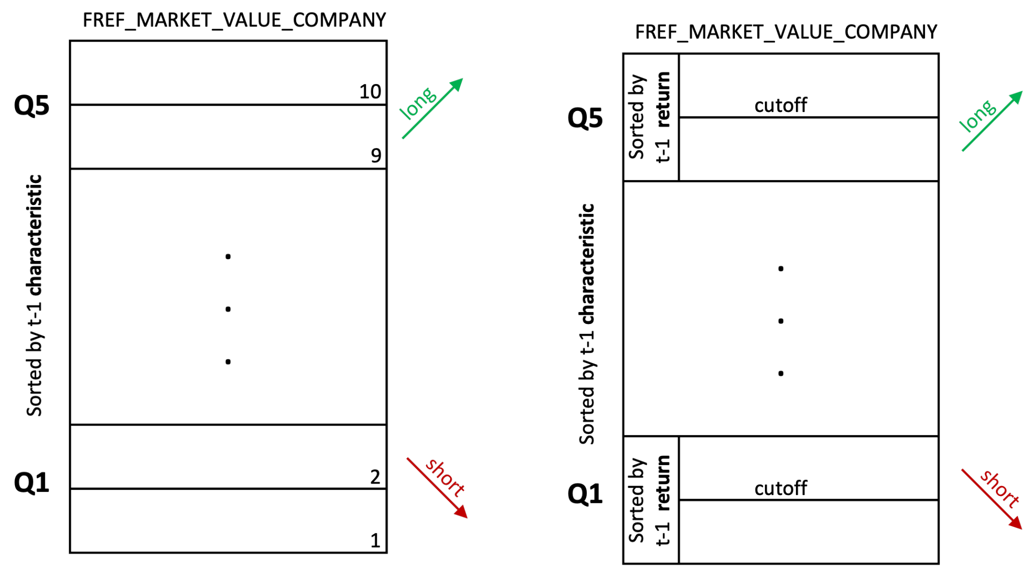
For each of the mean reversion strategies, we first determine via their underlying factor momentum version which factors in our sample performed well and which performed badly. For MRCSFM that means comparing recent factor returns over the cross-section of the factor sample, encoding those with above-median returns as *factor momentum winners* and those with below-median returns as *factor momentum losers*. For MRTSFM, we determine for each factor separately, whether it belongs to the factor momentum winners or losers, depending on its recent return being above or below zero. After this pre-selection has occurred, we zoom in on each factor separately. The core concept of momentum claims that securities that have been sorted into the *winner* portfolio will continue to perform well, whereas those in the *loser* portfolio will continue their downtrend [Jegadeesh and Titman, 1993]. Our assumption goes one step further and hypothesizes that, for factor momentum winners, the worst performing winner securities (lower half of return-sorted Q5 quantile) will show a reversion to the quantile mean return, effectively ending up in the upper half of the Q5 quantile in the next timestep, whereas the best performing loser securities (upper half of return-sorted Q1 quantile) will sink to the bottom half of the Q1 quantile, respectively. The spreads between these securities then make up the mean reversion winner portfolio. For factor momentum losers, we expect the opposite. The final mean reversion strategy then trades the spread between the returns of its winner and loser portfolio.

[26] cf. also McLean and Pontiff [2016] for the detrimental effects of academic research on anomaly payoffs.

The attentive reader may have noticed that the baseline of our proposed strategy resembles a 9-2 strategy which builds factors not from five quantiles but 10 deciles, trading the spread between the 9^{th} and 2^{nd} decile for each factor. A core difference, however, is that after sorting securities by their respective *factor characteristic* to form the factor quantiles, we sort securities within the Q5 and Q1 quantiles again, this time after their *lagged one-month performance* before splitting the quantiles further into two halves each. This does not alter the factor return as the constituents of the quantiles stay the same, it is just their order within the quantiles that changes. Figure 4.1 tries to illustrate this difference between the strategies.

Figure 4.1: Comparison of a 9-2 strategy and our approach

(a) 9-2 Strategy (b) Our Approach



4.2 Plain Vanilla Version

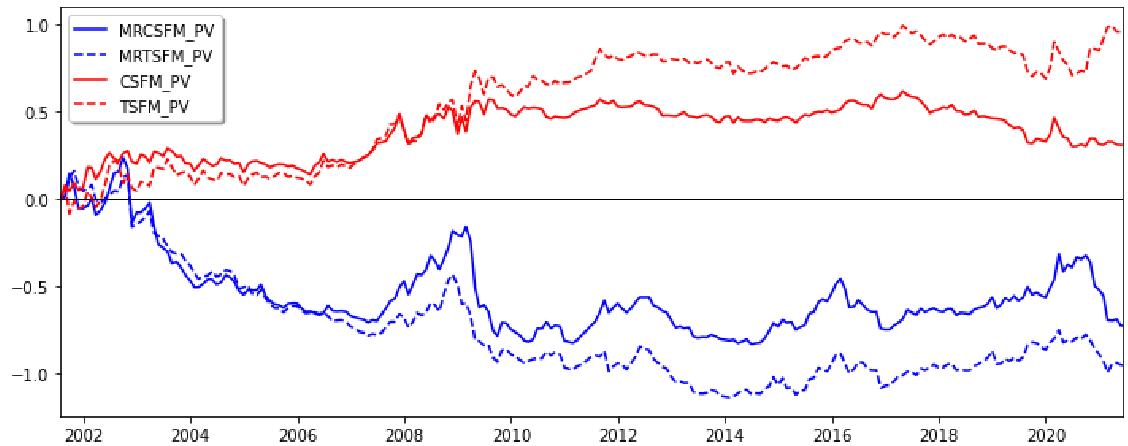
Just like in sections 3.3 and 3.4, we first define a *plain vanilla* (PV) version of MRCSFM and MRTSFM that follows above description, using CSFM_PV and TSFM_PV (one month formation period, one month holding period, cutoffs at the median/zero) versions to determine factor momentum winners and losers. We introduce yet another cutoff parameter that determines the level at which factor Q5 and Q1 quantiles are split and leave it at 0.5 for now, splitting the quantiles into equal halves after their constituents have been sorted by their past-month return. Figure 4.2 illustrates the cumulative log returns of MRCSFM_PV and MRTSFM_PV, as well as their respective factor momentum versions.

Both, MRCSFM_PV and MRTSFM_PV, run negatively over the whole timeframe. Although not statistically significant, annual expected returns are negative (-3.62% and -4.75%, t-values of 0.92 and 1.50) with moderately high standard deviations (17.60% and 14.14%), leading to small but negative Sharpe ratios of -0.21 and -0.34, respectively. The Sortino ratios of -0.22 and -0.37 lead to the conclusion that most of the volatility within MRCSFM_PV and MRTSFM_PV's returns is downside risk. During the financial crisis of 2008/09 both mean reversion strategies seem to perform well at first, but then give back all (and more) they have made during this brief window in the subsequent months. The most basic version of factor momentum based mean reversion fails to meet our expectations, leaving us to belief that this effect does not exist, at least within our dataset limited in time and investment universe.

4.3 Grid Search

Although our results from section 4.2 are not encouraging, we are curious to find out whether different parameter values lead to scenarios in which we can document robust factor momentum based mean reversion. We therefore construct a grid search environment that tests all possible hyperparameter combinations. For the underlying factor momentum dimension, we allow the same values for formation periods (1, 3, 6, 9 and 12 months) and cutoff values (0.1, 0.2, 0.3, 0.4 and 0.5 for CSFM; 0, 0.005, 0.01 and 0.015 for TSFM) as in the sensitivity analyses of sections 3.3 and 3.4.

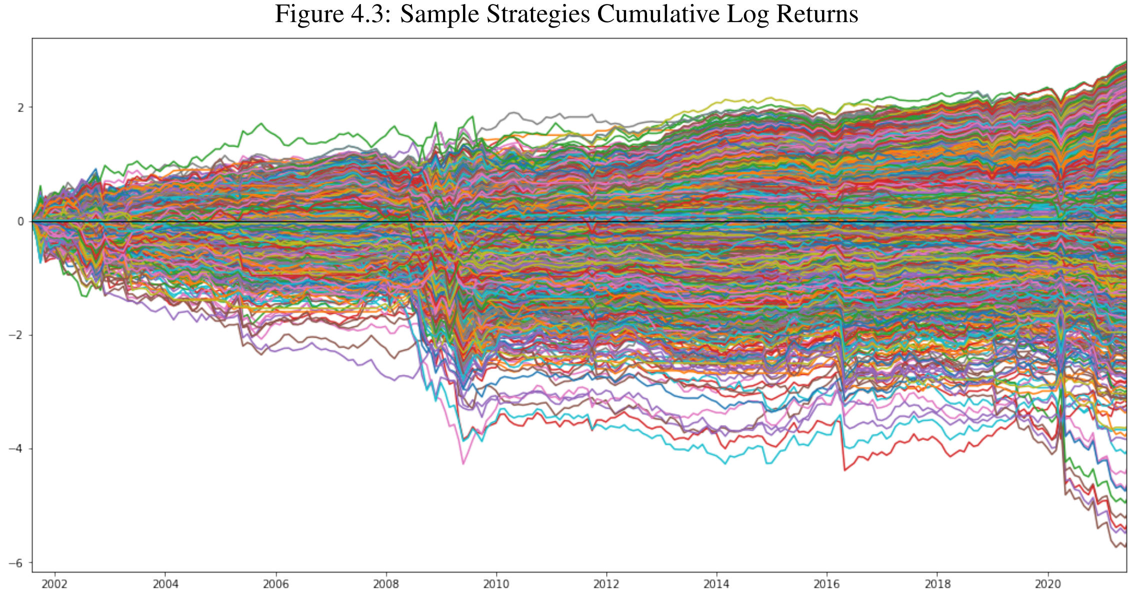
Figure 4.2: MR_PV Cumulative Log Returns



For the mean reversion dimension, we allow cutoff values to range between 0.1 (split return-sorted quantile into ten deciles and observe the securities in the 10^{th} and 1^{st} decile), 0.2, 0.3, 0.4 and 0.5 (base case quantile halving) to allow for more extreme security selections. Furthermore, we define seventeen additional, slight permutations of the plain vanilla strategy, each trading the spread of the quantile constituents according to the respective cutoff value. For thoroughness sake, we also consider non-spread trading long/short-only variants, fully aware that these do not represent the 'original' idea of mean reversion anymore. If we find that these permutations outperform the properly defined mean reversion strategies, we can reject our main assumption with more confidence. In addition to the final, spread trading mean reversion strategies, we also consider the respective winner and loser portfolios separately. Lastly, we also allow 9-2 variants as briefly described in section 4.1, i.e. not sorting by return within the quantiles. Therefore, half of the MRCSFM and MRTSFM strategy permutations actually describe factor momentum substrategies while the other half tries different settings for factor momentum based mean reversion.

As a result, we are presented with a total of 24,300 strategy permutations. Figure 4.3 visualizes the cumulative log returns of a randomly-drawn subsample of 5,000 strategy permutations.

Figure 4.3: Sample Strategies Cumulative Log Returns



It becomes apparent that most permutations run negatively. In the complete sample, the average expected return is 1.51% at an average standard deviation of 17.35%. Sharpe and Sortino ratios are, on average, 0.09 and 0.12.

High level clustering[27] did not reveal any meaningful patterns or scenarios in which MRCSFM or MRTSFM perform reliably well or significantly different from the factor momentum strategies. This further strengthens our result from above, that the mean reversion effect we have hoped to find does not exist in our data set. Table 4.1 compares average annual statistics for those strategy permutations that sort securities by their lagged one-month return within the top and bottom quantiles before being split by the respective cutoff value (*sorted by return*), and those that follow a 9-2 strategy and leave the order in the quantiles as is (*sorted by factor characteristic*).

Table 4.1: Average Statistics for Quantiles sorted by return/factor characteristic

	E[R]	STD	t-value	Sharpe	Range(Sharpe)
sorted by return	1.12	19.36	1.08	0.07	1.40
sorted by factor characteristic	1.89	15.33	1.33	0.11	1.66

Rounded to two decimals. E[R] and STD are reported in percent.

The difference in the reported metrics is small, the standard deviations are large, esp. in contrast to the average expected annual returns, and nothing turns out to be statistically significant. The ranges between the smallest and highest Sharpe ratios are quite large at 1.40 and 1.66, hinting at at least some well-performing strategy permutations. These, however, can be attributed rather to a lucky combination of parameters, highly *overfitted* to the observed timeframe, than any robust anomaly effect. This comparison shows that there is no significant difference between our mean reversion approach and a 9-2 strategy.

Going forward, we will consider two more variations of MRCSFM and MRTSFM each, in addition to above plain vanilla versions. MRCSFM_BS and MRTSFM_BS describe the *best Sharpe* (BS) permutations of the grid search that fulfill the requirements for a proper factor momentum based mean reversion trading strategy, i.e. splitting the top and bottom quantiles according to the cutoff value after sorting their respective constituents by their lagged one-month returns, going long in the bottom half of Q5 and short in the top half of Q1. For a permutation to be considered it must also trade the spread between its winner and loser portfolio. If we apply all of these requirements to the grid search result space, expected returns range from -16.17% to -2.96% (t-values from 0.55 to 2.56, 1.54 on average) for MRCSFM and from -12.69% to -0.93% (t-values from 0.30 to 2.36, 1.35 on average) for MRTSFM. It appears that no matter which parameter values are chosen for the underlying factor momentum signal generation, as well as for the cutoff on the mean reversion level, each strategy permutation runs negatively and is statistically unsignificant on average. The best Sharpe versions merely manage to achieve annual expected returns of -3.77% and -0.96% (t-values of 0.55 and 0.30), at standard deviations of 30.85% and 14.48%, leading to Sharpe ratios of -0.12 and -0.07, respectively.

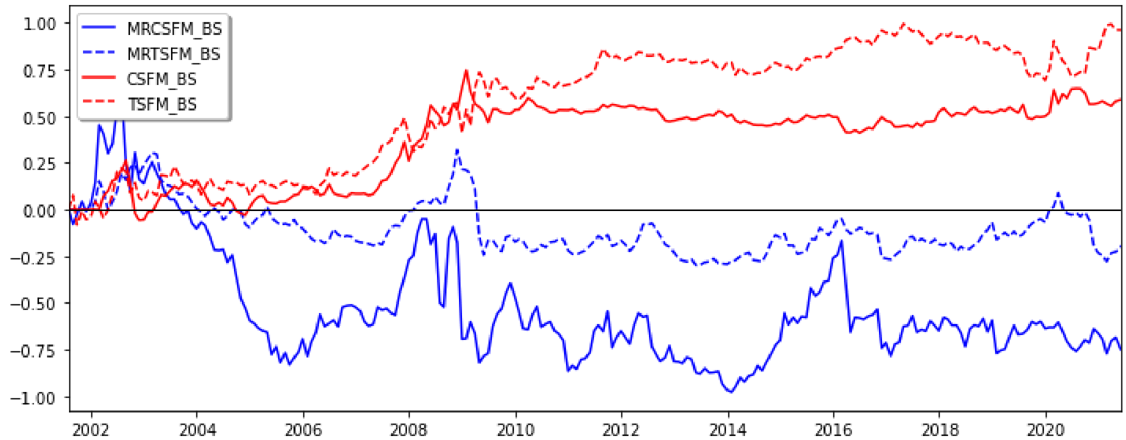
Figure 4.4 illustrates the cumulative log returns of MRCSFM_BS and MRTSFM_BS, as well as those of the best Sharpe versions of CSFM and TSFM. Note, that now, the best Sharpe versions of CSFM and TSFM are not necessarily those that lead to the best Sharpe versions of MRCSFM and MRTSFM.

In fact, MRCSFM_BS builds upon a one-month formation period with a cutoff level of 0.1 at the factor momentum level, as well as a cutoff level of 0.1 at the mean reversion level. Factors are hence considered winners/losers if their lagged one-month return falls into the top/bottom 10% of the cross-section of all factors in the sample. Securities within the top and bottom factor quantiles are first sorted by their past month's return and then split into the best/worst 10%. For the winner portfolio, the strategy therefore goes long in the worst 10% of the top quantile and short, respectively, in the best 10% of the bottom quantile of all factor momentum winners. The loser portfolio is formed the same way and the final MRCSFM_BS strategy then trades the spread between the winner and loser portfolio.

MRTSFM_BS on the other hand considers six months for the formation period with a cutoff level of zero at the factor momentum level, as well as a cutoff level of 0.4 at the mean reversion level. Factors are hence considered winners/losers if their average returns over the past six months were positive/negative. Securities in the top and bottom quantiles are first sorted by their lagged one-month returns and then split into the best/worst 40%. For the winner portfolio, the strategy therefore goes long in the worst 40% of the top quantile and short, respectively, in the best 40% of the bottom quantile for all factor momentum winners. The loser portfolio is formed the same way and the final MRTSFM_BS strategy then trades the spread between the winner and loser portfolio.

[27]See Figure A.5.

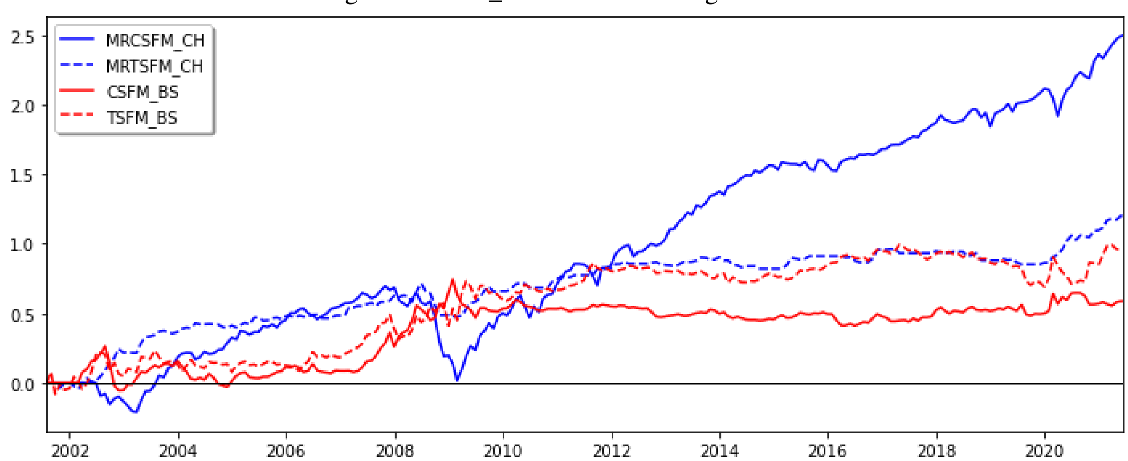
Figure 4.4: MR_BS Cumulative Log Returns



We conclude from both, the base cases and the results of our exhaustive sensitivity analysis, that factor momentum based mean reversion does not exist as a reliable and profitable effect in our underlying data set. Nevertheless, to showcase the maximum potential of MRCSFM and MRTSFM, we yet again pick best Sharpe versions from the grid search result space, but this time with no restrictions at all, i.e. these new *cheat* (CH) strategies do not represent pure factor momentum based mean reversion strategies anymore, at least as we defined them above.

Figure 4.5 illustrates the cumulative log returns of these highly overfitted MRCSFM_CH and MRTSFM_CH strategies.

Figure 4.5: MR_CH Cumulative Log Returns



Whereas MRTSFM_CH barely outperforms its factor momentum benchmark (TSFM_BS) at average annual returns of 6.02% (t-value of 3.45) and a standard deviation of 7.80% (Sharpe ratio of 0.77), MRCSFM_CH stands out with a stellar performance of 12.51% in expected returns (t-value of 3.68) at 15.21% standard deviation (Sharpe ratio of 0.82). In fact, MRCSFM_CH closely resembles a scaled version of the S&P 500 index itself, suggesting that it identifies the main performance drivers within the index' investment universe and trades them with leveraged positions.[28] Apparently, the further we stray from the original idea, the better the results become. Nevertheless, it is possible to construct attractive trading strategies with the right parameters, running the risk of possibly overfitting to the backtesting period. Since these versions are not representative of the pure factor momentum based mean reversion strategy (they do not necessarily trade spreads or are made up of classic long/short legs), we will not go into further detail about their specifications and will not consider them anymore going forward.

[28] See Figure A.6 for an illustrative comparison of MRCSFM_CH and MRTSFM_CH's cumulative log returns to the S&P 500.

4.4 Mean Reversion Statistics

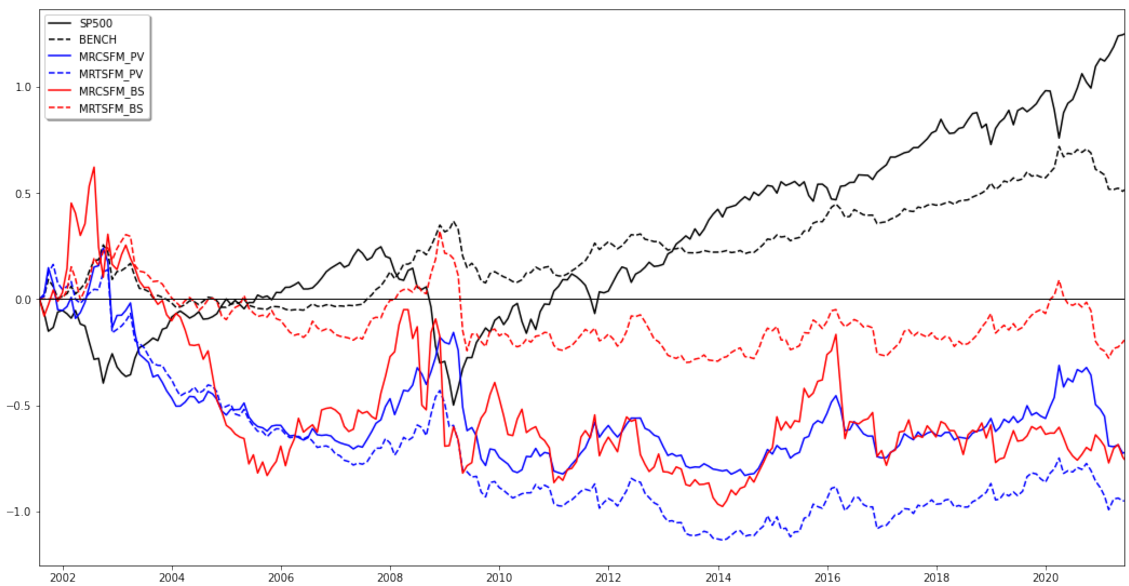
Table 4.2 explores the correlation between the four definitive mean reversion strategies. Both plain vanilla versions are somewhat closely correlated at 0.84, whereas the best Sharpe versions remain only moderately correlated (0.48). MRCSFM_BS is rather weakly correlated to the plain vanilla versions (0.36 and 0.47), whereas MRTSFM_BS is rather closely correlated (0.76 and 0.71).

Table 4.2: Mean Reversion Return Correlations

	MRCSFM_PV	MRTSFM_PV	MRCSFM_BS	MRTSFM_BS
MRCSFM_PV	1.00			
MRTSFM_PV	0.84	1.00		
MRCSFM_BS	0.36	0.47	1.00	
MRTSFM_BS	0.76	0.71	0.48	1.00

Figure 4.6 finally visualizes the cumulative log returns of the mean reversion strategies against the naïve benchmark BENCH and the S&P 500.

Figure 4.6: Mean Reversion, Cumulative Log Returns



Neither of the mean reversion strategies manages to outperform the benchmarks. In fact, even a simple, equal-weight buy and hold strategy of every factor in the sample performs significantly better than the lavish mean reversion strategies. Although they perform comparatively well in times of crises where the S&P 500 experienced large drops (esp. 2008/09 but also during 03/2020), they cannot manage to keep up this performance for long. Taken together with the fact that all of the mean reversion strategies are negatively correlated to the S&P 500 (-0.20 to -0.61), they could potentially serve as some sort of tail-hedge strategy that triggers in extreme downturn market phases. But even in such three or more sigma events, CSFM or TSFM strategies would probably still be a better choice (cf. Figure 3.8). Whereas the S&P 500's most extreme monthly returns were at -16.94% and 12.68%, monthly returns of the mean reversion strategies generally show heavier tails (with extreme returns ranging from -40.34% to -22.89% on the downside and 11.64% to 43.91% on the upside).[29] In the next section, we will thus explore different ex-post portfolio combinations to test whether the mean reversion strategies can be combined together in useful ways with factor momentum and other popular factors.

Table 4.3 serves as a reminder of the expected annual returns, standard deviations, Sharpe and Sortino ratios and Figure 4.7 additionally shows corresponding error terms and t-values for these metrics.

[29]See Figure A.7 for the return distributions of the mean reversion strategies.

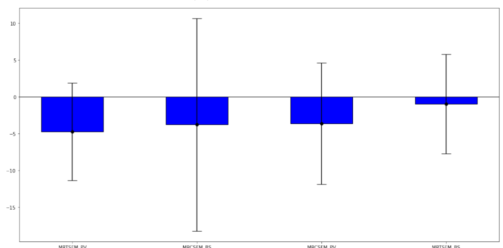
Table 4.3: Mean Reversion Statistics

Factor	E[R]	STD	Sharpe	Sortino
MRCSFM_PV	-3.62	17.60	-0.21	-0.22
MRTSFM_PV	-4.75	14.14	-0.34	-0.37
MRCSFM_BS	-3.77	30.85	-0.12	-0.14
MRTSFM_BS	-0.96	14.48	-0.07	-0.08

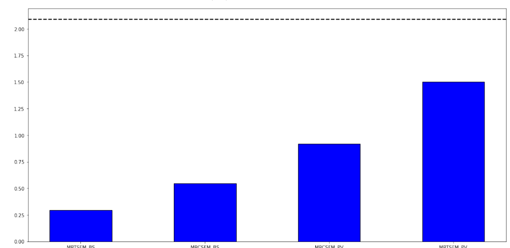
Rounded to two decimals. Superscript of * denotes statistical significance at p = 0.05 level. E[R] and STD are reported in percent.

Figure 4.7: Mean Reversion Statistics

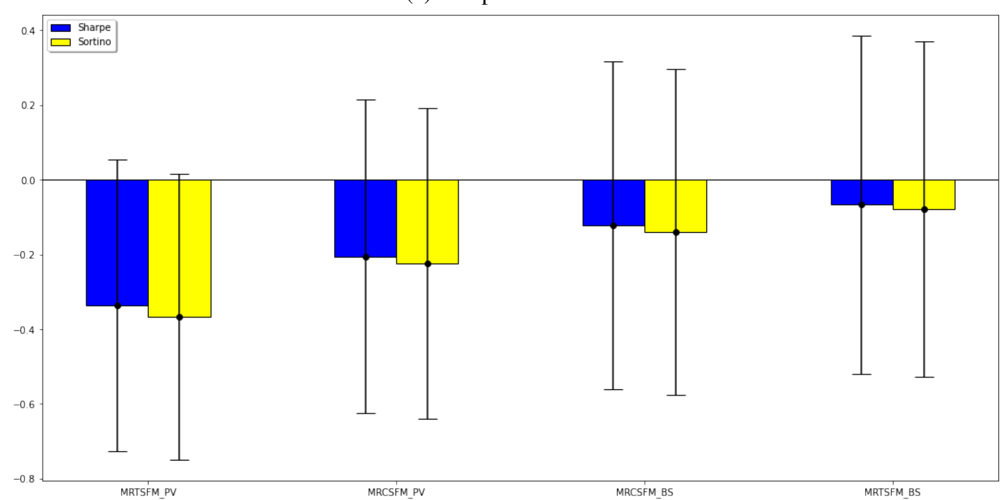
(a) Raw E[R]



(b) t-values



(c) Sharpe & Sortino



4.5 Regressions

To conclude this section, we regress each of the factor momentum strategies (CSFM_PV, TSFM_PV, CSFM_BS, TSFM_BS) and each of the mean reversion strategies (MRCSFM_PV, MRTSFM_PV, MRCSFM_BS, MRTSFM_BS) over the S&P 500 (Table 4.4), Fama and French [2015]'s five factors, as well as a 10 factor model building upon the five factor model and expanding it by the popular factors UMD, LTR, STR, BAB and QMJ (Table 4.5).[30]

None of the intercepts (alphas) over the S&P 500 are significant at 95% confidence. We can, however, see a clear trend with the factor momentum strategies all sharing positive signs for their risk-adjusted returns, whereas the mean reversion strategies mostly underperforming the index.

For the FF5 and FF10 alphas we can observe a very similar effect with the factor momentum strategies earning mostly small positive risk-adjusted returns (although statistically unsignificant) and the mean reversion strategies

[30] cf. Carhart [1997], Bondt and Thaler [1985], Jegadeesh [1990], Frazzini and Pedersen [2014], Asness et al. [2017]. UMD, LTR and STR monthly returns are fetched from the Kenneth R. French website, QMJ and BAB monthly returns from AQR's website (`https://www.aqr.com/Insights/Datasets?&page=2#filtered-list`).

Table 4.4: Regression Results over S&P 500

	Alpha	Error	t-value	R^2_{adj}
CSFM_PV	1.74	5.07	0.68	-0.26
TSFM_PV	4.61	6.72	1.35	-0.33
CSFM_BS	2.91	5.17	1.11	-0.42
TSFM_BS	4.61	6.72	1.35	-0.33
MRCSFM_PV	0.72	6.32	0.23	34.54
MRTSFM_PV	-1.94	5.53	0.69	22.27
MRCSFM_BS	-0.58	3.28	0.35	10.86
MRTSFM_BS	-4.94	9.48	1.03	18.24

Rounded to two decimals. Alpha, Error and R^2_{adj} are reported in percent.

Table 4.5: Regression Results

	over FF5				over FF10			
	Alpha	Error	t-value	R^2_{adj}	Alpha	Error	t-value	R^2_{adj}
CSFM_PV	0.04	5.26	0.02	0.33	-1.40	5.47	0.50	0.19
TSFM_PV	3.29	7.03	0.92	-0.74	0.97	7.22	0.26	1.39
CSFM_BS	1.16	5.41	0.42	-1.29	1.25	5.62	0.44	-1.38
TSFM_BS	3.29	7.03	0.92	-0.74	0.97	7.22	0.26	1.39
MRCSFM_PV	-4.56	5.05	1.78	61.56	-5.60	3.59	3.07	81.92
MRTSFM_PV	-5.47	4.96	2.17	42.62	-5.26	4.62	2.24	53.71
MRCSFM_BS	-2.04	3.31	1.21	16.64	-2.96	3.28	1.78	24.34
MRTSFM_BS	-7.91	9.63	1.62	22.41	-7.85	8.88	1.74	38.81

Rounded to two decimals. Alpha, Error and R^2_{adj} are reported in percent.

underperforming the factor models (statistically signficant for MRTSFM_PV over FF5 and FF10, and MRCSFM_PV over FF10).

In the regression tests of Ehsani and Linnainmaa [2019], Gupta and Kelly [2018], CSFM underperforms TSFM, i.e. when TSFM is regressed over CSFM, its alpha is positive and statistically significant, but not vice versa. In our case, CSFM_PV regressed over TSFM_PV results in an alpha of -1.54% (t-value of 1.15), whereas for the regression of TSFM_PV over CSFM_PV, we get an alpha of 3.05% (t-value of 1.72). Although not significant at 95% confidence, we observe a very similar effect as the authors. For the mean reversion strategies, we can observe an opposite effect with MRCSFM_PV regressed over MRTSFM_PV having an alpha of 1.37% (t-value of 0.65) and -2.29% (t-value of 1.35) vice versa. The best Sharpe versions of both, factor momentum and mean reversion, do not show such effects.

5 Portfolio Optimization

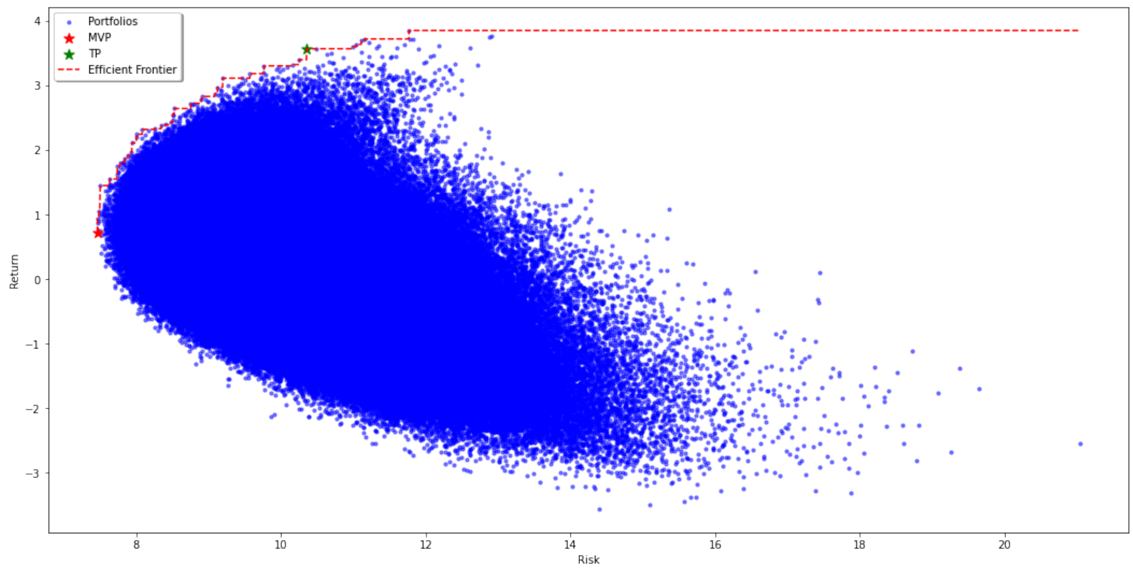
The past section has demonstrated that factor momentum based mean reversion does not depict a robust and profitable trading strategy on its own. Nevertheless, due to its tail-hedge character, it could still potentially contribute in a meaningful way to a broader portfolio. This section therefore tests various ex-post portfolio combinations in the spirit of Markowitz [1952, 1987] of the various mean reversion and factor momentum strategies, as well as the same popular factors as in section 4.5.

5.1 Efficient Frontier, Minimum Variance & Tangency Portfolio (long-only)

We use a Monte Carlo Simulation with 1,000,000 iterations to form an equal number of portfolios assigning random weights to each of the passed factor momentum and mean reversion strategies. Weights within each portfolio have to sum up to one and positions are long-only. Figure 5.1 illustrates each of these portfolios via its risk-return dimensions.

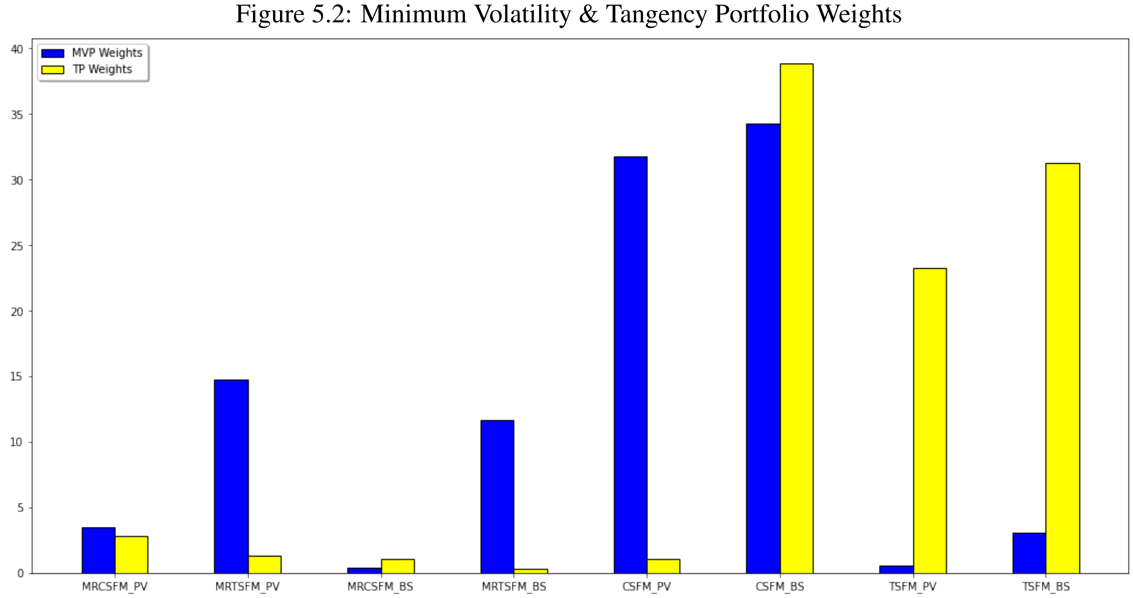
The majority of the simulated portfolios falls in the lower half of the parabola which can be traced back to the mostly negative performance of the mean reversion strategies. We plot the *Efficient Frontier* in red which marks those portfolios

Figure 5.1: Monte Carlo Portfolios & Efficient Frontier



with the highest expected returns for a given level of risk (maximum Sharpe ratio). The *Minimum Variance Portfolio* (MVP; marked by a red star in above figure) achieves an expected annual return of 0.72% at a standard deviation of 7.46% (Sharpe ratio of 0.10). The *Tangency Portfolio* (TP; marked by a green star) returns, on average, 3.57% per year at a standard deviation of 10.35% (Sharpe ratio of 0.34). Figure 5.2 reports the corresponding weights of the considered strategies for the MVP (blue) and TP (yellow).

Figure 5.2: Minimum Volatility & Tangency Portfolio Weights



The MVP aims at minimizing the portfolio risk, hence it naturally places its largest weights in those strategies with the lowest standard deviations (CSFM_BS, CSFM_PV, MRTSFM_PV and MRTSFM_BS with standard deviations of 11.63%, 11,41%, 14.14% and 14.48%) and tries to avoid those with large volatility (e.g. MRCSFM_BS and TSFM_PV/TSFM_BS at standard deviations of 30.85% and 15.12%). The TP on the other hand aims at maximizing the Sharpe ratio and therefore places large emphasis on those strategies with high return/risk ratios such as CSFM_BS, TSFM_PV/TSFM_BS (Sharpe ratios of 0.25 and 0.32).

5.2 Optimal Portfolio (long-only)

The results from the Monte Carlo Simulation already give a good approximation for the full-sample mean-variance optimized portfolio. To corroborate our results from above, we utilize a linear optimizer that, rather than finding a tangency portfolio by random trial and error, takes the expected returns and covariance matrix of the supplied strategies as well as a set of constraints (e.g. lower/upper bounds for strategy weights) and a function to minimize as inputs to calculate the optimal portfolio. For the long-only variant, we do not allow leverage, i.e. weights $\in [0, 1]$ and need to sum up to one. Note that by minimizing the negative ratio of return and risk, we actually solve for the strategy weights that maximize the portfolio Sharpe ratio. Table 5.1 summarizes the results over various portfolio combinations and reports the respective portfolio Sharpe ratios.[31] We use cash with zero mean and zero standard deviation as a proxy for a riskfree investment.

Table 5.1: Ex-post Optimal Portfolio Weights & Sharpe Ratios (long-only)

		Tangency Portfolio Weights				
Strategy	Indiv. Sharpe	MR	FM	MR + FM (a)	a + FF10 (b)	b + CASH
MRCSFM_PV	-0.21	0.00		0.00	0.00	0.00
MRTSFM_PV	-0.34	0.00		0.00	0.00	0.00
MRCSFM_BS	-0.12	0.00		0.00	0.00	0.00
MRTSFM_BS	-0.07	1.00		0.00	0.00	0.00
CSFM_PV	0.14		0.00	0.00	0.00	0.00
CSFM_BS	0.25		0.47	0.47	0.04	0.04
TSFM_PV	0.32		0.26	0.26	0.02	0.02
TSFM_BS	0.32		0.26	0.26	0.02	0.02
Mkt-RF	0.49				0.22	0.21
SMB	0.26				0.08	0.08
HML	-0.14				0.00	0.00
RMW	0.47				0.00	0.00
CMA	0.20				0.14	0.14
UMD	0.00				0.00	0.00
STR	0.24				0.02	0.02
LTR	-0.10				0.00	0.00
BAB	0.70*				0.13	0.13
QMJ	0.36				0.34	0.34
CASH	0.00					0.00
Sharpe		-0.07	0.40	0.40	1.54*	1.54*

Rounded to two decimals. Superscript of * denotes statistical significance at p = 0.05 level.

The optimizer tries to avoid the mean reversion strategies as much as possible, setting all of their weights to zero as long as there are other strategies to construct the portfolio from. Otherwise, it chooses and invests 100% in the strategy with the best Sharpe ratio (MRTSFM_BS). Taking into account that both versions of TSFM are the same, an optimal portfolio of all factor momentum strategies invests roughly equally into CSFM and TSFM, logically choosing the best Sharpe versions over the plain vanilla versions. When combined with the 10 popular risk factors, it becomes evident that neither the mean reversion strategies nor the factor momentum strategies can contribute to a broad investment portfolio in a meaningful way.[32] In the most exhaustive portfolio configurations, QMJ, the market portfolio and CMA make up the three biggest positions with an aggregated weight of roughly 70%. It is interesting to note that UMD's (individual stock momentum) weighting is also set to zero due to the factor's bad performance over the observed timeframe (expected annual returns of 0.01% at a standard deviation of 17.47%, resulting in a Sharpe ratio of practically 0).

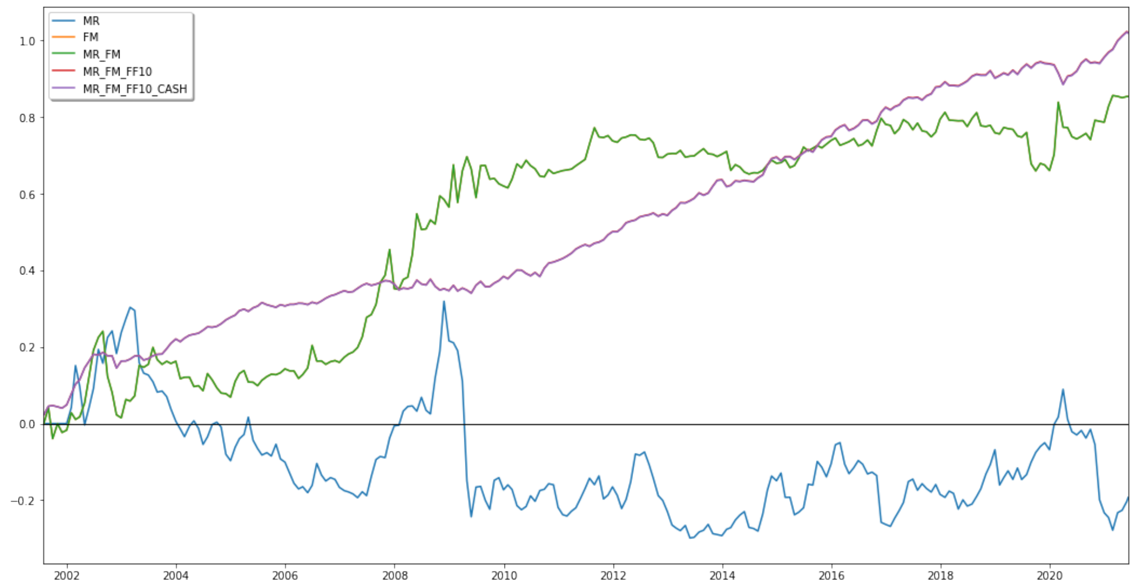
Figure 5.3 illustrates the cumulative log returns of the five optimal portfolios. Since the portfolio that trades all mean reversion and factor momentum strategies applies the same weights as the portfolio trading only the factor momentum

[31] As already noted throughout section 3.4, TSFM_PV and TSFM_BS actually describe the same strategy but are kept separate to keep the consistency to the other strategy variants.

[32] Note, however, that the 10 risk factors are calculated on a different, much larger and more diverse investment universe than our S&P 500 based factor momentum and mean reversion strategies.

strategies, its cumulative log return is the same and hence overlaid in the figure. We observe the same for the two large optimal portfolios since the weight in cash is negligible (0.16%).

Figure 5.3: Long-Only Optimal Portfolios Cumulative Log Returns



5.3 Optimal Portfolio (long-short)

We repeat the linear optimization from above, this time allowing long and short positions with some leverage (weights $\in [-2, 2]$). Portfolio weights, however, still have to sum up to one and short selling is allowed with no costs. Table 5.2 summarizes the results.

The optimizer makes heavy use of the added short selling option to bet against the worst performing strategies (MRTSFM_PV, CSFM_PV and MRCSFM_PV) and finance leveraged long positions in the most promising ones (MRTSFM_BS, TSFM_PV/TSFM_BS). A portfolio consisting of all mean reversion strategies now achieves a positive Sharpe ratio of 0.08 despite every strategy having a negative Sharpe ratio individually. For the mean reversion and factor momentum optimal portfolios, however, final Sharpe ratios hardly increase in comparison to those of the long-only portfolios. In the more exhaustive portfolios, our strategies and the riskfree investment mainly serve as financiers for long positions in the popular risk factors. QMJ, CMA and the market portfolio remain amongst the most strongly weighted positions.

Figure 5.4 illustrates the cumulative log returns of the five optimal portfolios.

At Quantumrock (and other quantitative hedge funds such as e.g. Renaissance Technologies [Zuckerman, 2019]), it is common practice to not only trade one definite version of a trading strategy, but rather build mean-variance optimized portfolios of all permutations whose backtest return series are statistically significant. Referring back to the grid search result space in section 4.3, 4,600 out of the 24,300 strategy permutations for MRCSFM and MRTSFM showed returns different from zero at 95% confidence. It could have been interesting to explore the profitability of ex-post or even rolling window tangency portfolio versions of these significant strategies but that was beyond the scope of this project study.[33]

6 Robustness & Implementability

The past sections have thoroughly analyzed the proposed mean reversion strategies as well as their underlying, signal generating, cross-sectional and time-series factor momentum strategies. In hedge fund practice, however, statistics

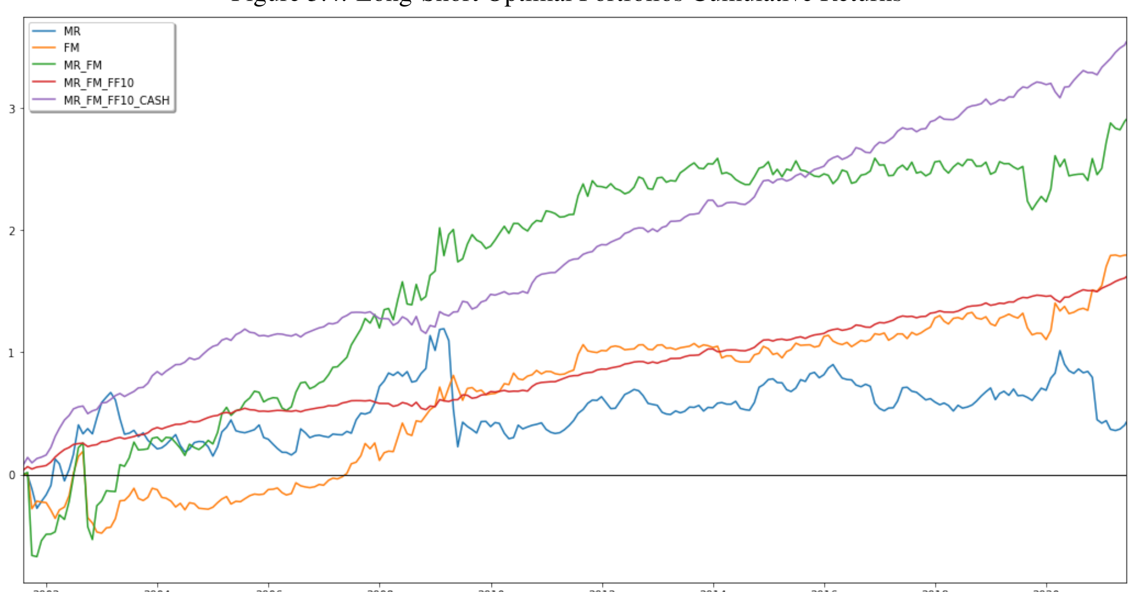
[33]. . . and of our laptops' computing power.

Table 5.2: Ex-post Optimal Portfolio Weights & Sharpe Ratios (long-short)

		Tangency Portfolio Weights				
Strategy	Indiv. Sharpe	MR	FM	MR + FM (a)	a + FF10 (b)	b + CASH
MRCSFM_PV	-0.21	0.96		0.10	-0.24	-0.53
MRTSFM_PV	-0.34	-2.00		-2.00	-0.01	-0.02
MRCSFM_BS	-0.12	0.04		-0.10	-0.01	-0.03
MRTSFM_BS	-0.07	2.00		1.41	0.09	0.20
CSFM_PV	0.14		-2.00	-1.72	-0.11	-0.24
CSFM_BS	0.25		0.85	1.10	0.05	0.12
TSFM_PV	0.32		1.07	1.10	0.05	0.11
TSFM_BS	0.32		1.07	1.10	0.05	0.11
Mkt-RF	0.49				0.22	0.49
SMB	0.26				0.11	0.25
HML	-0.14				-0.24	-0.54
RMW	0.47				0.16	0.35
CMA	0.20				0.31	0.69
UMD	0.00				-0.00	-0.00
STR	0.24				0.02	0.05
LTR	-0.10				-0.11	-0.23
BAB	0.70*				0.21	0.48
QMJ	0.36				0.45	1.00
CASH	0.00					-1.23
Sharpe		0.08	0.40	0.41	1.73*	1.71*

Rounded to two decimals. Superscript of * denotes statistical significance at p = 0.05 level.

Figure 5.4: Long-Short Optimal Portfolios Cumulative Returns



such as expected returns, t-values and Sharpe ratios are not the only metrics of interest. Thus, this section briefly explores various other robustness measures, loosely inspired by Pedersen [2015], as well as goes over practicability of the strategies and their implementability into an existing trading environment. Lastly, after the example of Gupta and Kelly [2018], we will take a closer look at the turnover and the implications of trading costs onto our strategies.

6.1 Key Metrics

Table 6.1 reports expected annual returns, volatilities (standard deviations) and Sharpe ratios for each of our main strategies. In addition to these already known metrics, we point out each strategy's maximum drawdown (Max. DD), expected annual excess return (E[ER]), ex-post tracking error (TE) and information ratio (IR) relative to the S&P 500, as well as their monthly 95% Value at Risk (VaR) and expected Shortfall (CVaR).

Table 6.1: Key Robustness Metrics

Factor	E[R]	STD	Sharpe	Max. DD	E[ER]	TE	IR	VaR	CVaR
MRCSFM_PV	-3.62	17.60	-0.21	-65.75	-13.73	29.38	-0.47	9.24	13.32
MRTSFM_PV	-4.75	14.14	-0.34	-72.68	-14.30	25.17	-0.57	7.03	10.60
MRCSFM_BS	-3.77	30.85	-0.12	-79.77	-13.16	36.82	-0.36	13.21	21.39
MRTSFM_BS	-0.96	14.48	-0.07	-46.10	-10.64	26.00	-0.41	5.88	10.49
CSFM_PV	1.55	11.41	0.14	-27.12	-6.96	18.87	-0.37	4.55	6.86
TSFM_PV	4.80	15.12	0.32	-26.14	-3.56	20.70	-0.17	5.56	9.04
CSFM_BS	2.94	11.63	0.25	-28.53	-5.50	18.66	-0.29	4.43	7.76
TSFM_BS	4.80	15.12	0.32	-26.14	-3.56	20.70	-0.17	5.56	9.04

Rounded to two decimals. Superscript of * denotes statistical significance at p = 0.05 level. E[R], STD, Max. DD, E[ER], TE, VaR and CVaR are reported in percent.

Maximum drawdowns for the mean reversion strategies are, on average, at -66%, and for the factor momentum strategies around -27% which demonstrates how badly negative expected returns and high standard deviations of the mean reversion strategies combine. The best Sharpe version of MRCSFM even documents a maximum drawdown of -80%, requiring a 400% gain to revert back to the initial investment value - a killing blow to any professional investment strategy. The expected excess returns show that none of the selected strategies manages to outperform the S&P 500.[34] Tracking errors over the whole timeframe are relatively high and the information ratios (Sharpe ratio on excess returns) are all negative - another sign for underperformance. Finally, the VaR describes the maximum expected monthly loss with 95% confidence which centeres around 9% for the mean reversion strategies and around 5% for the factor momentum strategies. For tail loss events, i.e. those cases where losses surpass the VaR expectations, CVaR values average to 14% and 8%, respectively.

In Figure 6.1[35], we go into more detail of strategy evaluation on the example of MRTSFM_BS as the overall best performing mean reversion strategy.

Panels (a) and (b) display MRTSFM_BS's volatility and Sharpe ratio on a 24-month rolling window and Panel (c) illustrates the high water mark (HWM) as the cumulative maximum return, as well as drawdowns in absolute and relative terms. The strategy performs mostly negative, but in extreme, black swan[36] events such as during the financial crisis of 2008/09 and the COVID pandemic market shock in 03/2020, it is able to generate positive returns. In the second half of the 2008/09 recession, however, MRTSFM_BS fails to capitalize on its prior performance and quickly reverts back to pre-crisis levels, quickly accumulating a large drawdown from which it fails to recover until the end of the timeframe. Its rolling volatility during that time spikes at around 8% and stands in no comparison to any other time period, not even the COVID crisis. Over the rest of the observed timeframe, the strategy does not manage to lift its HWM, which oftentimes serves as an important threshold for charging performance fees to investors [Pedersen, 2015], above historic levels. Panel (d) finally extrapolates the monthly VaR and CVaR estimations over the course of one whole year, showcasing the extreme potential (tail) losses.

[34] This partially contradicts the results from Table 4.4 where we find small but positive alphas for the factor momentum strategies regressed over the S&P 500 and smaller negative alphas for the mean reversion strategies. However, we can attribute these differences to the weak explanatory power of the linear regression model (R^2_{adj} practically zero for factor momentum and in the low deciles for mean reversion).

[35] NBER Recession data is taken from `https://fred.stlouisfed.org/series/USREC`.

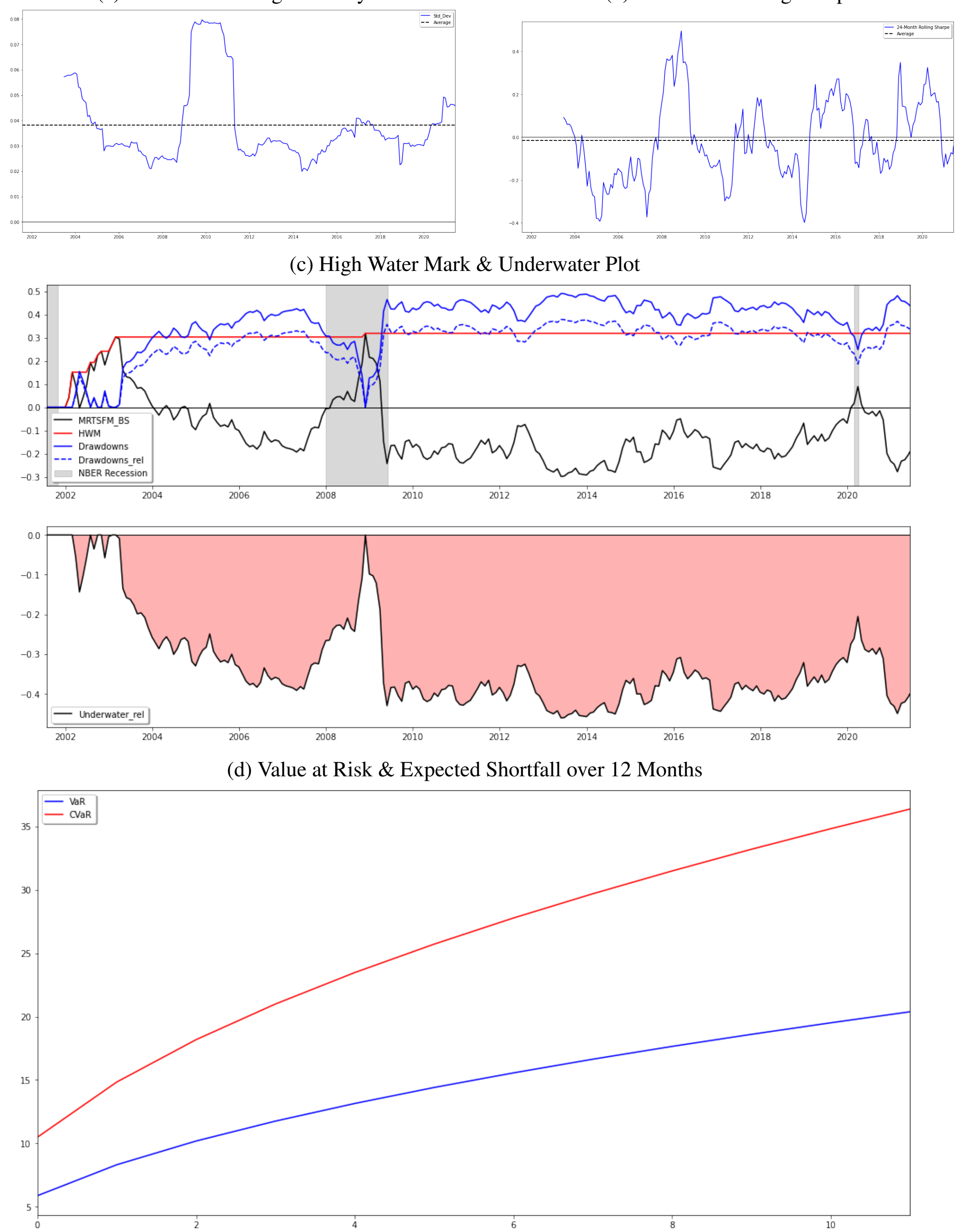
[36] cf. Taleb [2007].

Figure 6.1: Exemplary Robustness Metrics for MRTSFM_BS

(a) 24-Month Rolling Volatility (b) 24-Month Rolling Sharpe

(c) High Water Mark & Underwater Plot

(d) Value at Risk & Expected Shortfall over 12 Months



6.2 Ticker-based Backtest & Industry Groups

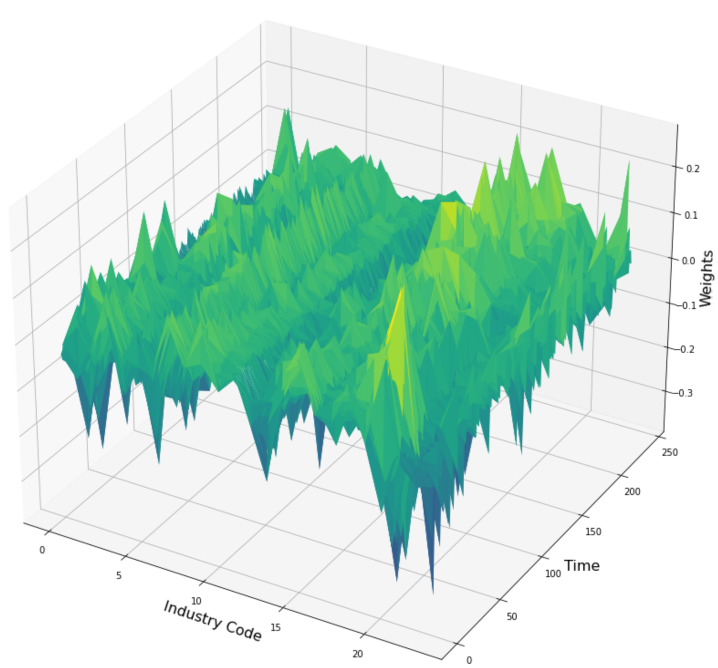
Constructing elaborate trading strategies on the factor level (and beyond) can quickly become very abstract and academic. In practice, however, trading algorithms need to work on the most basic, ticker, level to generate proper buy and sell orders.

For this reason, we built a backtesting environment that, rather than calculating the performance of the factor momentum and mean reversion strategies directly via the factor quantile spreads, traces the signals from the strategies all the way back to the individual ISINs representing the tickers within the S&P 500 investment universe. With the information of which tickers were traded at which timesteps with which weightings, we were then able to calculate the strategy performances by taking the respective realized ticker returns directly from our tensor (cf. section 2.5). Since the final ticker-based return series matched the factor-level return series for every strategy, we can ensure the correctness and practical realizability of our proposed trading strategies.

During this process, we needed to calculate a portfolio weight matrix for each trading strategy that features all tradable securities on one axis and each timestep on the other. Cell values are then filled with the corresponding weights for each ticker such that we can determine the exact portfolio composition at every timestep. Figure A.8 illustrates such a portfolio weight matrix for the MRTSFM_PV strategy. Green values encode long positions and red values stand for short positions. The more intense the cell color, the higher the respective weight within the ticker, and vice versa.

Our original data was divided by industry group, which, together with the portfolio weight matrices, allowed us to calculate the industry exposure for each of our trading strategies. Figure 6.2 illustrates exemplary for MRTSFM_PV how loadings across industries evolve over time. Upon closer inspection, it is possible to identify some troughs (primarily shorted industries such as *Energy & Natural Resources*, *Insurance* and *Banking*) and spikes (primarily long positions in certain industries such as *Food & Beverages*, *Real Estate* and *Consumer Services*). Here, weights do not necessarily sum up to one since some tickers are assigned to several industries.

Figure 6.2: Exemplary Industry Group Weighting (MRTSFM_PV)

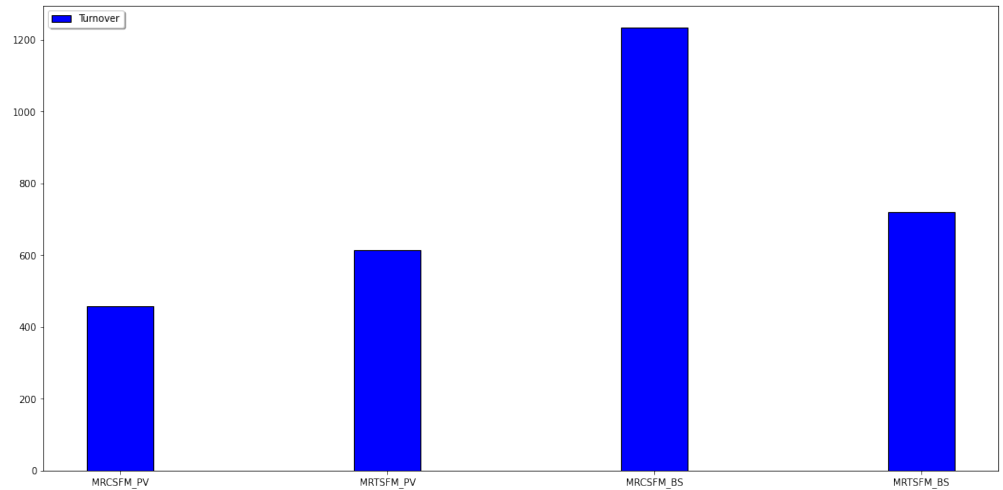


6.3 Turnover & Transaction Costs

Calculating said portfolio weight matrices also came with the advantage of determining the absolute changes in ticker weights over every timestep whose sum will be referred to as *turnover* in the following. Panel (a) of Figure 6.3 illustrates the total turnover for each of the mean reversion trading strategies over the whole timeframe of 20 years.

Figure 6.3: Turnover and Net Sharpe Ratios of Mean Reversion Strategies

(a) Turnover



(b) Sharpe Ratios, before and after cost

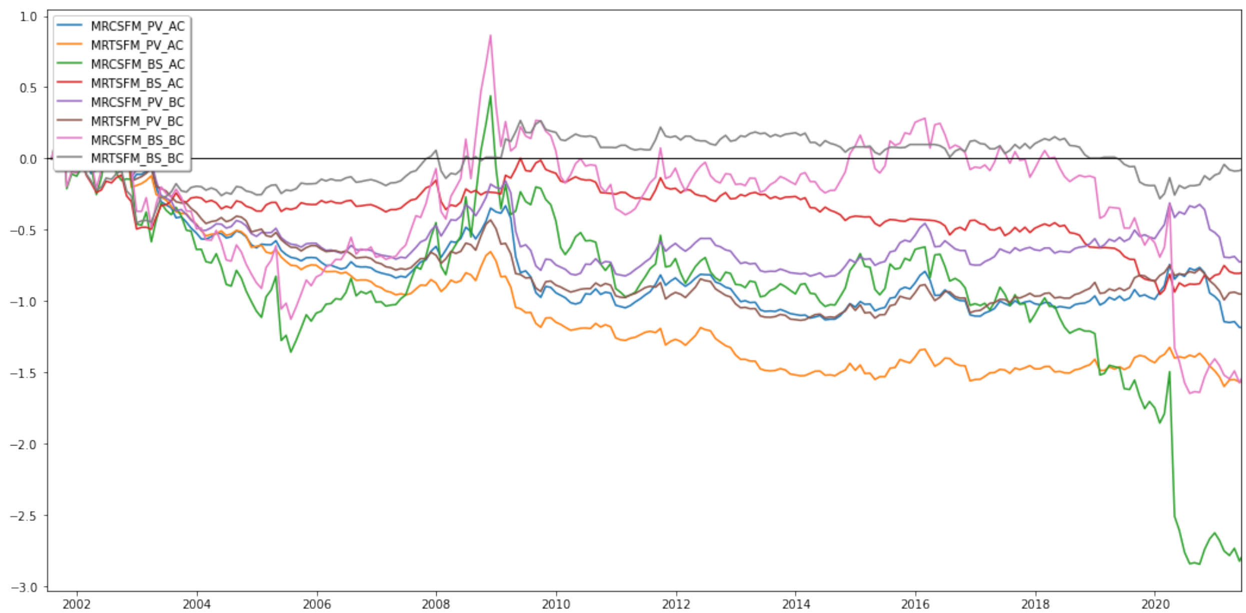
MRCSFM_BS involves by far the most trading out of all mean reversion strategies. This should come as no surprise since its underlying factor momentum winners and loser are determined as the top/bottom 10% of the cross-section of all factors in the sample (cf. section 4.3). Additionally, MRCSFM_BS uses an extreme cutoff of 0.1 to split factor quantiles after sorting securities by their lagged one-month return. These extreme selection criteria aim at identifying the tickers with the highest potential for mean reversion in the next timestep, but clearly come at the cost of high portfolio turnover. Its plain vanilla counterpart which determines well and badly performing factors over the cross-sectional median of returns and splits the return sorted quantiles into equal halves, on the other hand, trades only around a third as much. For MRTSFM_BS and MRTSFM_PV, we can observe a similar effect, with the more sophisticated security selection of the best Sharpe version requiring more trading than the plain vanilla strategy. In the end, it is a tradeoff between selecting potentially more promising securities and higher turnover, at least for monthly portfolio rebalancing.

Following Gupta and Kelly [2018], Frazzini and Pedersen [2014], we assume incurred costs of 10 basis points per unit of turnover to calculate Sharpe ratios net of transaction costs. Panel (b) of Figure 6.3 compares gross and net Sharpe ratios for the different mean reversion strategies.

The large drops from gross to net Sharpe ratios suggest that, even if the mean reversion strategies were able to generate positive returns, the involved trading would probably wipe out any profits entirely, further discouraging their usage in practice. It is interesting to note that although MRCSFM_BS has the highest turnover, its relative change from before to after cost Sharpe ratio is not as large as one would expect. We are not entirely sure what exactly causes this anomaly, but our educated guess is that the high volatility associated with MRCSFM_BS (30.85%) sometimes achieves positive returns high enough to offset a large portion of the trading costs.[37]

Figure 6.4 finally illustrates gross and net cumulative log returns for the factor momentum based mean reversion strategies.

Figure 6.4: Gross and Net Cumulative Log Returns for Mean Reversion Strategies



7 Limitations & Conclusion

The S&P 500 is undoubtedly a very efficient and liquid market environment of predominantly large and highly profitable companies, making any efforts to beat it especially difficult. In this paper, we have shown that even active trading strategies making use of sophisticated concepts such as momentum on the factor level, as well as mean reversion, cannot reliably outperform the market. There are, however, some limitations to our results.

Our dataset is very limited, especially when compared to those of other works in this area before us, taking only S&P 500 index constituents (survivorship bias-free, but instead biased regarding size, liquidity and growth) over the last 20 years into account.[38] We work with a sample of only 10 factors, many of which are highly correlated since they are calculated from the same pricing information, resulting in merely 2,400 monthly return data points to build upon and limiting the predictability and significance of our results. Furthermore, we argue that markets in the post 2000 era have become much more efficient which further complicates finding undiscovered anomalies.

A factor's past return being informative about its future return is a key assumption underlying the idea of factor momentum. The majority of our factor returns, however, cannot be differentiated from zero on usual statistical

[37]The three highest monthly returns of the MRCSFM_BS strategy are 43.91%, 36.10% and 22.48%.

[38]Avramov et al. [2017], Gupta and Kelly [2018], Ehsani and Linnainmaa [2019], Arnott et al. [2021] are all considering the entirety of the U.S. securities market, including small caps.

confidence levels which might explain why we are also not able to find significant autocorrelation. This calls the ability of our sample to time factors into question.

Nevertheless, although yet again statistically unsignificant, our simple versions of CSFM and TSFM manage to come close to the performance of the S&P 500 within the observed timeframe, trading upside volatility for more robustness in returns with no large drawdowns and especially good performance during market turmoils. In our sensitivity analyses, we do not find any meaningful patterns in parameter combinations that lead to reliable and robust results. Future research could profit from formulating more sophisticated versions of CSFM and TSFM, e.g. as proposed by Gupta and Kelly [2018]. However, optimizing too many parameters based on past performance also increases the risk of overfitting to the sample data and bad generalization.

Performance of the factor momentum strategies is mainly driven by their short-legs which, due to short-sell restrictions, is oftentimes difficult to implement in practice, imposing yet another limitation onto our results.

We document negative performance for the factor momentum based mean reversion strategies. Even a large scale grid search on various parameters does not result in reliably well performing strategy permutations. When regressed over the S&P 500, as well as five and 10 factor models, the plain vanilla as well as best Sharpe versions of MRCSFM and MRTSFM exclusively achieve (partially signficant) negative alphas. Even in broader portfolios of several trading strategies and popular factors, the mean reversion strategies are either weighted at zero or predominantly shorted to finance leveraged positions in the other options. Further robustness tests revealed that the mean reversion strategies are prone to large drawdowns and high volatilities, leading to negative excess returns over the market. Lastly, picking factor momentum constituents after their potential to revert to the quantile mean involves high amounts of trading which, even in the case of positive returns, would further decrease their profitability and hence attractiveness to industry practitioners.

Taken together, these results leave us to conclude that mean reversion on a factor momentum basis does not exist as a robust and profitable anomaly, at least in our limited dataset.

We encourage future research to work at the very least with more (diverse) factors and potentially even a larger investment universe, as well as longer time-series of data. During our sensitivity analyses, we kept the holding period constant at one month. A higher dimensional grid search environment with more sophisticated measures to extract potentially meaningful patterns (e.g. unsupervised machine learning based clustering) could lead to new insights as well. Avramov et al. [2017] use time-series predictive regressions on various lagged variables to forecast factor returns and utilize them as signals for their strategies rather than conditioning on past returns. Apart from introducing this interesting alternative to factor momentum signal generation, the authors even document higher risk-adjusted Fama and French [1993] three factor alphas for the predictive return strategies than for those built from past return signals. We think it could be of interest to further explore this approach, e.g. via the use of autoregressive models such as ARIMA/ARFIMA, or even novel machine learning techniques.

Convolutional Neural Networks (CNNs) are traditionally used in image recognition tasks, but since they are specialized in extracting local patterns and finding dependencies, we are confident that they could also be used on properly formatted financial time-series data. If past returns can predict future performance, CNNs should theoretically be able to pick up and build upon these patterns. Going even one step further, Recurrent Neural Networks (RNNs) with e.g. a Long Short-Term Memory (LSTM) architecture, have longer 'memory' and can therefore draw connections and identify recurrent patterns over long periods of time, leading to potentially even better factor return forecasts. Besides predicting returns, such machine learning models could also be used in the portfolio construction process by optimizing strategy weights via their forecasted return volatilities.

However, training such neural networks properly requires a multitude of our modest data set[39], which is why we leave these exciting new approaches to future research.

[39] . . . and would go above and beyond the scope of this project study . . .

References

Robert D. Arnott, Mark Clements, Vitali Kalesnik, and Juhani T. Linnainmaa. Factor momentum. *SSN Working Paper*, 2021.

Clifford S. Asness, Tobias J. Moskowitz, and Lasse Heje Pedersen. Value and momentum everywhere. *The Journal of Finance*, 68(3):929–985, 2013.

Clifford S. Asness, Andrea Frazzini, and Lasse Heje Pedersen. Quality minus junk. *SSN Working Paper*, 2017.

Doron Avramov, Si Cheng, Amnon Schreiber, and Koby Shemer. Scaling up market anomalies. *The Journal of Investing*, 26:89–105, 2017.

Rolf W. Banz. The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9:3–18, 1981.

David Blitz and Matthias X. Hanauer. Settling the size matter. *SSN Working Paper*, 2020.

Werner F.M. De Bondt and Richard Thaler. Does the stock market overreact? *The Journal of Finance*, 40(3):793–805, 1985.

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

Kalok Chan, Allaudeen Hameed, and Wilson Tong. Profitability of momentum strategies in the international equity markets. *The Journal of Financial and Quantitative Analysis*, 35(2):153–172, 2000.

V. Chavez-Demoulin, Paul Embrechts, and Sylvain Sardy. Extreme-quantile tracking for financial time series. *Swiss Finance Institute Research Paper*, (11-27), 2011.

Tarun Chordia, Richard Roll, and Avanidhar Subrahmanyam. Recent trends in trading activity and market quality. *Journal of Financial Economics*, 101:243–263, 2011.

Tarun Chordia, Avanidhar Subrahmanyam, and Qing Tong. Have capital market anomalies attenuated in the recent era of high liquidity and trading activity? *Journal of Accounting and Economics*, 58:41–58, 2014.

John H. Cochrane. Presidential address: Discount rates. *The Journal of Finance*, 66(4):1047–1108, 2011.

D. A. Dickey and Wayne A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366):427–431, 1979.

Gregor Dorfleitner. Why the return notion matters. *International Journal of Theoretical and Applied Finance*, 6(1): 73–86, 2003.

Sina Ehsani and Juhani T. Linnainmaa. Factor momentum and the momentum factor. *NBER Working Paper*, 2019.

Eugene F. Fama. Random walks in stock market prices. *Financial Analysts Journal*, 21:55–59, 1965.

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

Eugene F. Fama and Kenneth R. French. The cross-section of expected stock returns. *The Journal of Finance*, 47(2): 427–465, 1992.

Eugene F. Fama and Kenneth R. French. Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33:3–56, 1993.

Eugene F. Fama and Kenneth R. French. A five-factor asset pricing model. *Journal of Financial Economics*, 116:1–22, 2015.

Guanhao Feng, Stefano Giglio, and Dacheng Xiu. Taming the factor zoo: A test of new factors. *The Journal of Finance*, 75(3):1327–1370, 2020.

Andrea Frazzini and Lasse Heje Pedersen. Betting against beta. *Journal of Financial Economics*, 111:1–25, 2014.

Tarun Gupta and Bryan T. Kelly. Factor momentum everywhere. *SSN Working Paper*, 2018.

Campbell R. Harvey, Yan Liu, and Heqing Zhu. ... and the cross-section of expected returns. *The Review of Financial Studies*, 29:5–68, 2016.

Antti Ilmanen, Ronen Israel, Rachel Lee, Tobias J. Moskowitz, and Ashwin Thapar. How do factor premia vary over time? a century of evidence. *Journal of Investment Management*, 19(4):15–57, 2021.

Narasimhan Jegadeesh. Evidence of predictable behavior of security returns. *The Journal of Finance*, 45(3):881–898, 1990.

Narasimhan Jegadeesh and Sheridan Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1):65–91, 1993.

John Maynard Keynes. *THE GENERAL THEORY OF EMPLOYMENT INTEREST AND MONEY*. Macmillan and Co., Limited, 1936.

Andrew W. Lo. The statistics of sharpe ratios. *Financial Analysts Journal*, 58, 2003.

Harry Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

Harry Markowitz. *Mean-variance analysis in portfolio choice and capital markets*. Oxford [u.a.] : Blackwell, 1987.

R. David McLean and Jeffrey Pontiff. Does academic research destroy stock return predictability? *The Journal of Finance*, 71(1):5–32, 2016.

Tobias J. Moskowitz and Mark Grinblatt. Do industries explain momentum? *The Journal of Finance*, 54(4):1249–1290, 1999.

Lasse Heje Pedersen. *Efficiently Inefficient*. Princeton University Press, 2015.

William F. Sharpe. Mutual fund performance. *The Journal of Business*, 39(1):119–138, 1966.

William F. Sharpe. Adjusting for risk in portfolio performance measurement. *The Journal of Portfolio Management*, 1: 29–34, 1975.

William F. Sharpe. The sharpe ratio. *The Journal of Portfolio Management*, 21:49–58, 1994.

Robert J. Shiller. Do stock prices move too much to be justified by subsequent changes in dividends? *The American Economic Review*, 71(3):421–436, 1981.

Frank A. Sortino and Lee N. Price. Performance measurement in a downside risk framework. *The Journal of Investing*, 3:59–64, 1994.

Frank A. Sortino and Robert van der Meer. Downside risk. *The Journal of Portfolio Management*, 17:27–31, 1991.

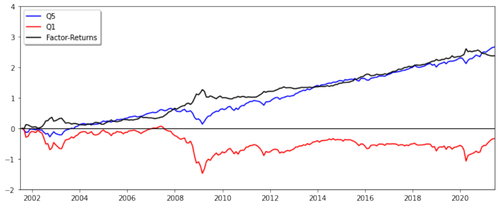
Nassim Nicholas Taleb. *The Black Swan: The Impact of the Highly Improbable*. Random House, 2007.

Gregory Zuckerman. *The Man Who Solved the Market: How Jim Simons Launched the Quant Revolution*. Portfolio; INT edition, 2019.

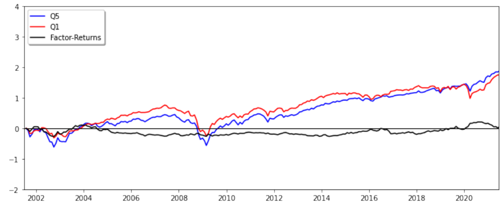
A Appendix

Figure A.1: Factor Quantile & Spread Portfolios

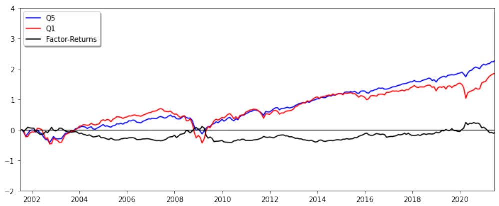
(a) FREF_MARKET_VALUE_COMPANY



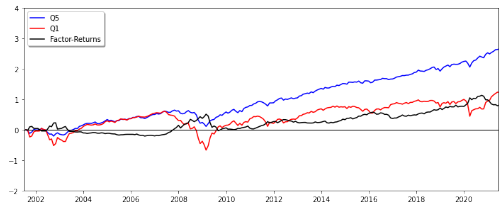
(b) FF_PE



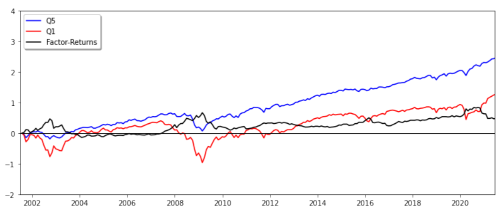
(c) FF_PSALES



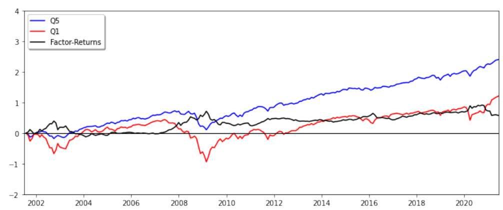
(d) FF_PBK



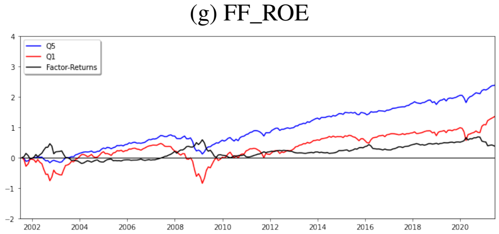
(e) FF_ROA



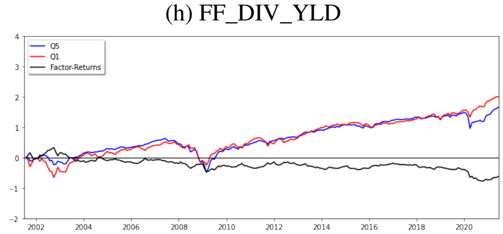
(f) FF_ROTC



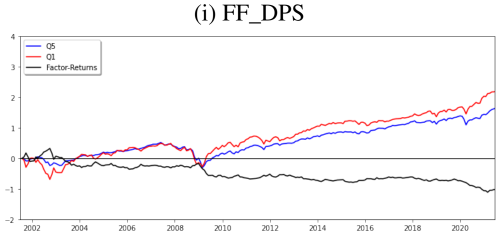
(g) FF_ROE



(h) FF_DIV_YLD



(i) FF_DPS



(j) FSI_DAYS_ANY_EXCHG

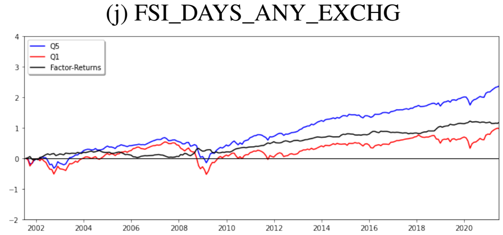
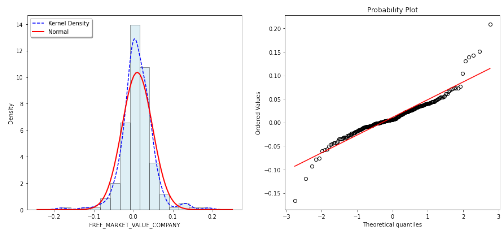
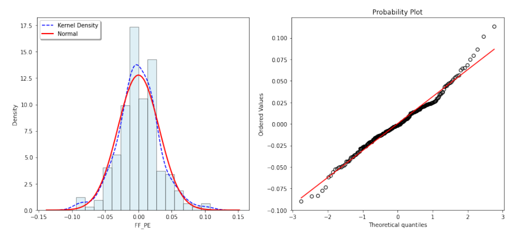


Figure A.2: Factor Returns Histograms & Q-Q Plots

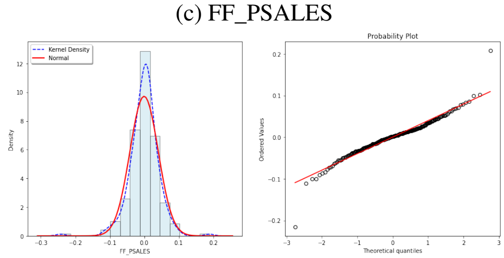
(a) FREF_MARKET_VALUE_COMPANY



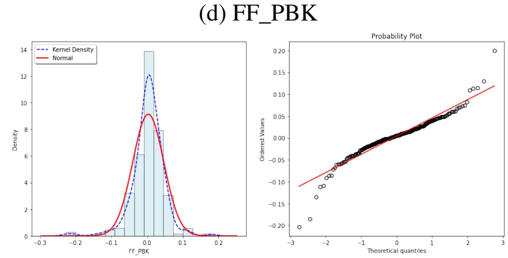
(b) FF_PE



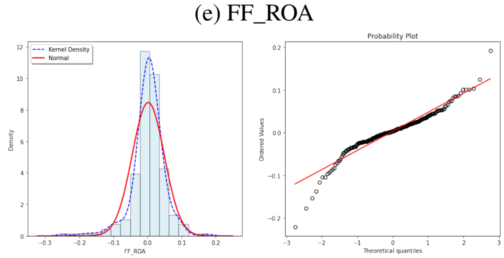
(c) FF_PSALES



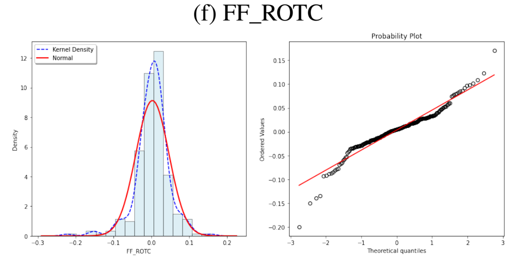
(d) FF_PBK



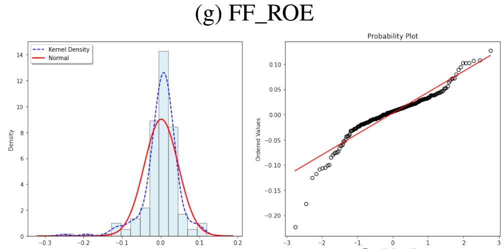
(e) FF_ROA



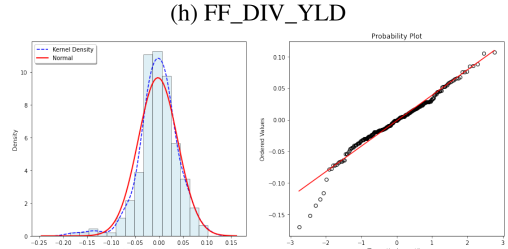
(f) FF_ROTC



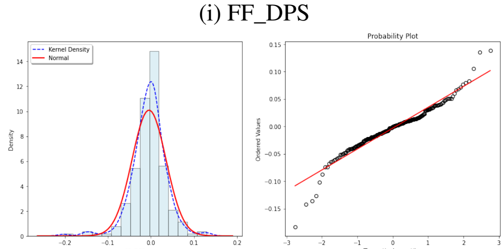
(g) FF_ROE



(h) FF_DIV_YLD



(i) FF_DPS



(j) FSI_DAYS_ANY_EXCHG

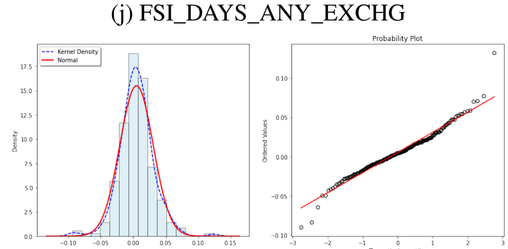


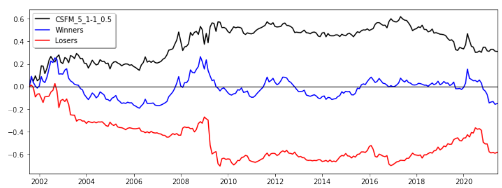
Table A.1: Augmented Dickey-Fuller Test Statistic

Factor	ADF t-value	AR(1) Coefficient
FREF_MARKET_VALUE_COMPANY	-14.09*	0.09
FF_PE	-8.06*	0.05
FF_PSALES	-6.21*	0.08
FF_PBK	-7.94*	0.12
FF_ROA	-8.20*	0.10
FF_ROTC	-8.38*	0.05
FF_ROE	-8.34*	0.13*
FF_DIV_YLD	-14.23*	0.08
FF_DPS	-13.75*	0.11
FSI_DAYS_ANY_EXCHG	-5.77*	0.10

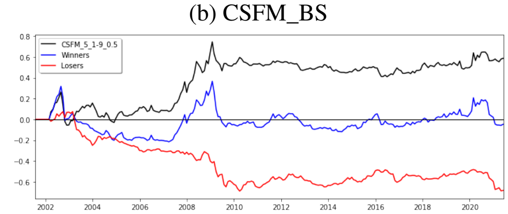
Rounded to two decimals. Superscript of * denotes statistical significance at p = 0.05 level.

Figure A.3: Factor Momentum Portfolios

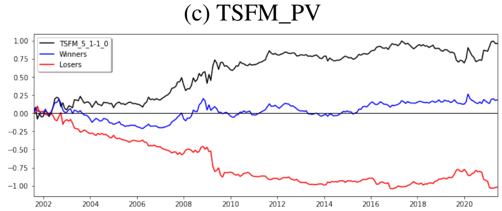
(a) CSFM_PV



(b) CSFM_BS



(c) TSFM_PV



(d) TSFM_BS

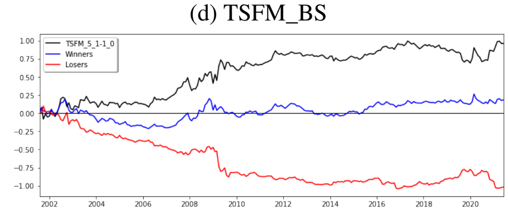


Table A.2: Average Annual Expected Returns & t-values for various Formation Periods and Cutoff Values

(a) CSFM

Cutoff	Formation Period					Formation Period				
Values	1	3	6	9	12	1	3	6	9	12
	Average E[R]					**t-values**				
0.1	2.38	-1.69	0.48	-0.44	-5.76	0.49	0.33	0.10	0.10	1.19
0.2	3.88	2.34	-0.24	2.25	-1.00	0.97	0.61	0.06	0.65	0.27
0.3	2.55	1.24	0.57	1.38	-0.75	0.74	0.36	0.17	0.42	0.24
0.4	1.78	1.48	-1.60	2.27	-1.09	0.62	0.50	0.51	0.80	0.37
0.5	1.55	1.42	-0.65	2.94	-0.07	0.61	0.55	0.24	1.13	0.03

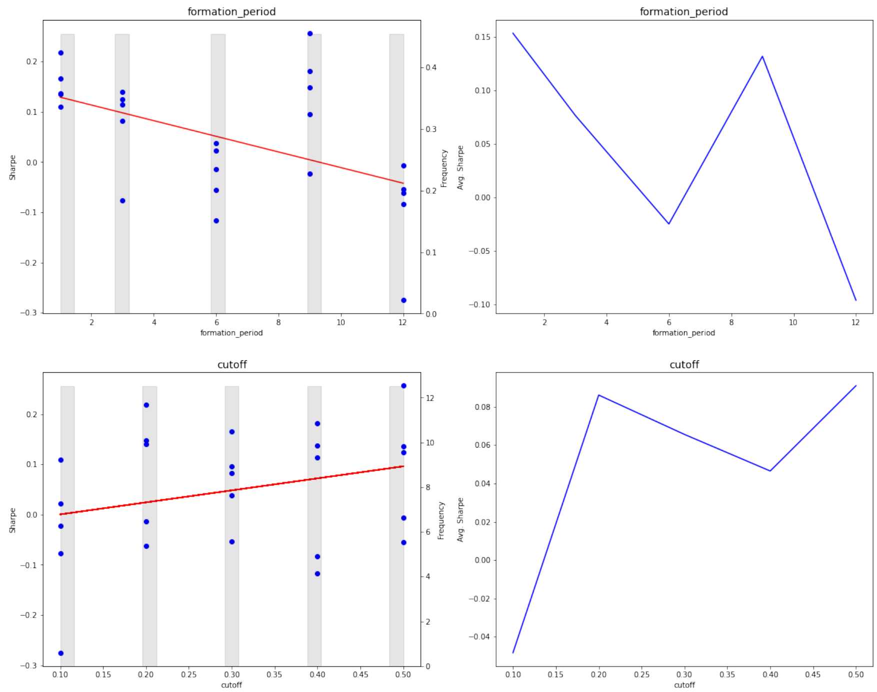
(b) TSFM

Cutoff	Formation Period					Formation Period				
Values	1	3	6	9	12	1	3	6	9	12
	Average E[R]					**t-values**				
0	4.80	2.23	2.25	3.87	0.60	1.42	0.71	0.67	1.19	0.20
0.005	2.72	2.67	2.36	2.53	0.03	0.75	0.78	0.64	0.72	0.01
0.010	3.09	3.28	0.01	2.69	-2.33	0.82	0.92	0.00	0.79	0.68
0.015	2.40	2.90	-0.78	0.74	-3.18	0.65	0.80	0.22	0.24	0.88

Rounded to two decimals. Average E[R] is reported in percent.

Figure A.4: Factor Momentum Features Performance & Frequency

(a) CSFM



(b) TSFM

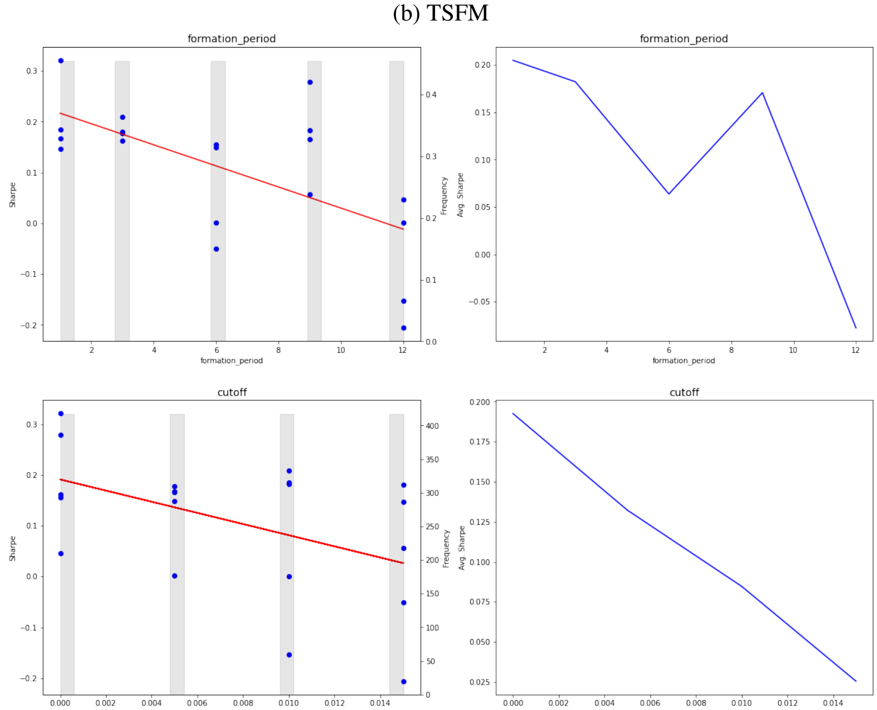
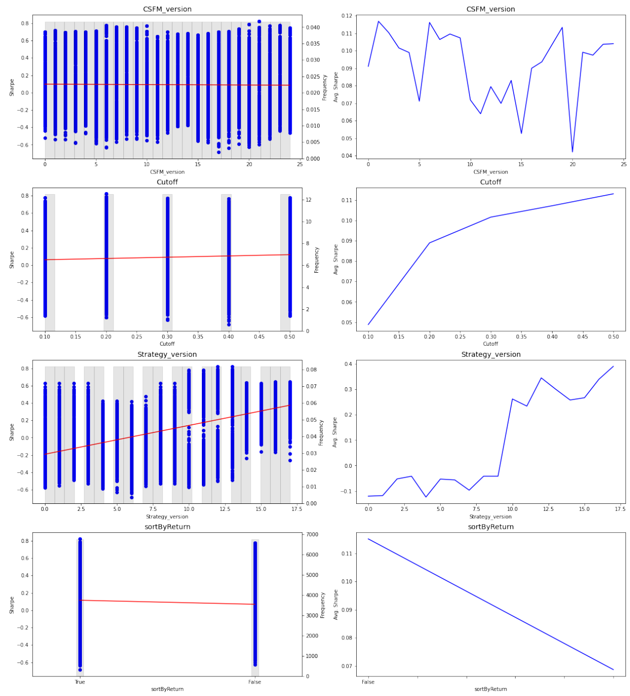


Figure A.5: Mean Reversion Features Performance & Frequency

(a) MRCSFM



(b) MRTSFM

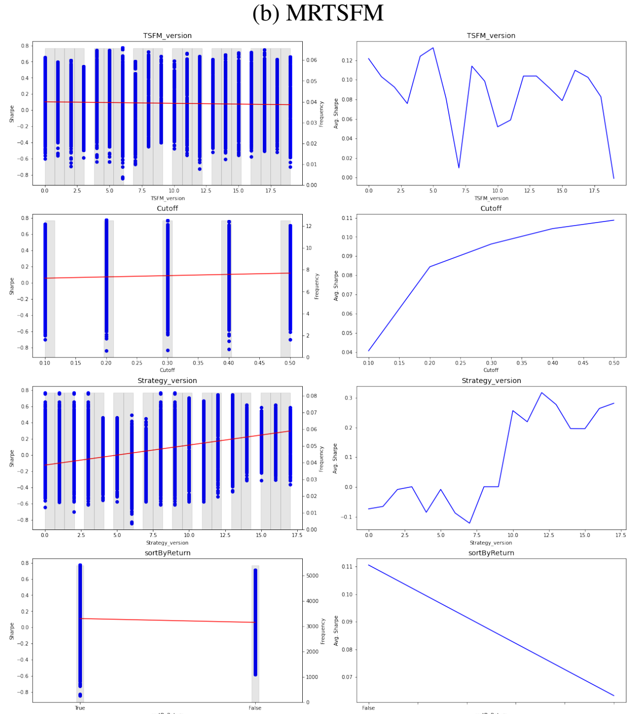


Figure A.6: MRCSFM_CH & MRTSFM_CH Cumulative Log Returns

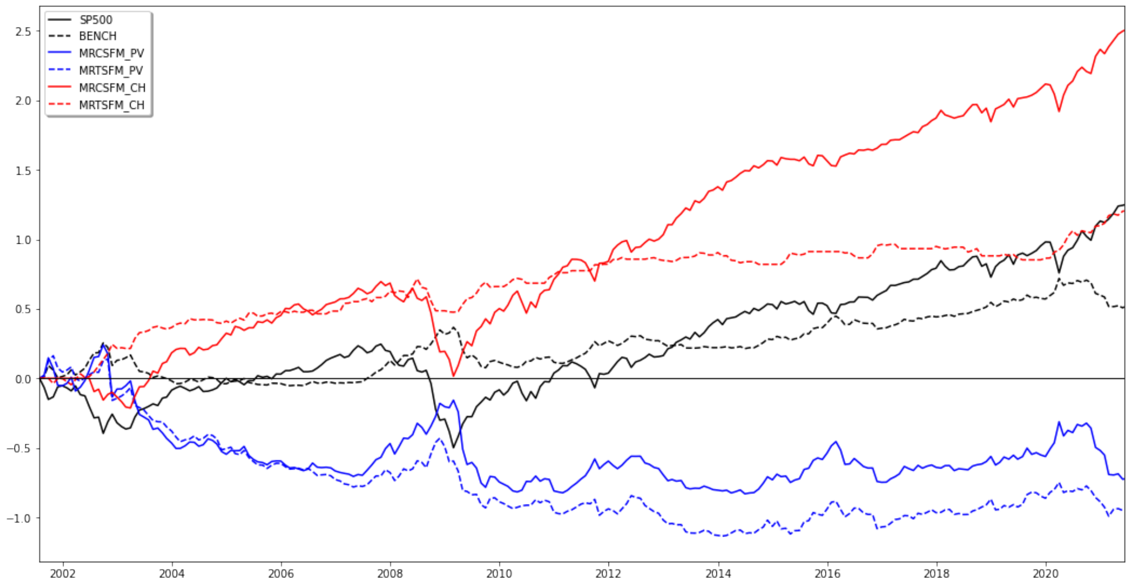
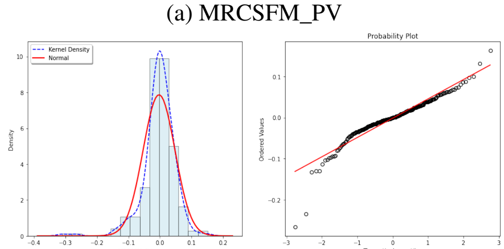
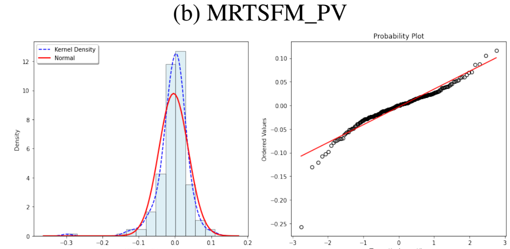


Figure A.7: Strategy Returns Histograms & Q-Q Plots

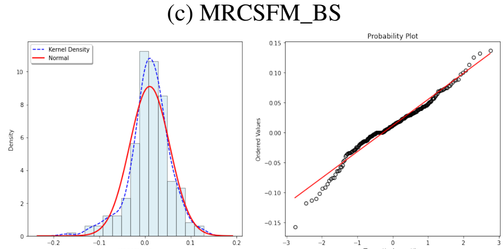
(a) MRCSFM_PV



(b) MRTSFM_PV



(c) MRCSFM_BS



(d) MRTSFM_BS

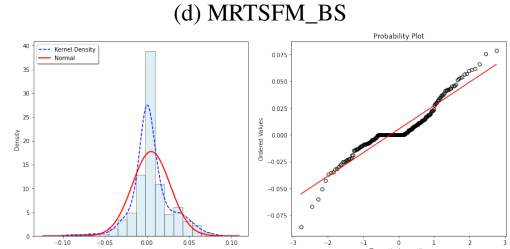


Figure A.8: Exemplary Ticker-based Backtest (MRTSFM_PV)