ERASMUS UNIVERSITY ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS
MSc Economics & Business
Master Specialisation Financial Economics

Multi-Factor Timing – A time series clustering approach

Author: Jakob Reithmann
Student number: 470544
Thesis supervisor: Dr. Esad Smajlbegovic
Finish date: May 4th, 2020

PREFACE AND ACKNOWLEDGEMENTS

I thank Dr. Smajlbegovic for his patience and support along this journey.

NON-PLAGIARISM STATEMENT

By submitting this thesis, the author declares to have written this thesis completely by himself, and not to have used sources or resources other than the ones mentioned. All sources used, quotes and citations that were literally taken from publications, or that were in close accordance with the meaning of those publications, are indicated as such.

ABSTRACT

Unsupervised learning tools in conjunction with macroeconomic forecasts can improve factor rotation strategies relative to traditional logit-based models and static equal-weight portfolios. Applying a popular subsequence time series clustering algorithm to identify common themes, or 'motifs', within six European factor return time series, I link these motifs to a variety of lagged macro-economic state variables using ordered logit regressions. The resulting, out-of-sample trading strategy generates strong, beta- and leverage-adjusted excess returns and substantially improved Sharpe, Omega and MAR ratios between 2011 and 2019, even after implementation costs. It also shows significant and exploitable alpha in Fama & French's six-factor model. A plain macro model simply calibrated on the factors' past standardized return time series, however, underperforms a static equal-weight portfolio and doesn't generate significant alphas. Overall, the findings hint at the importance of selectivity when calibrating logit or probit regression-based prediction models.

Keywords:

Unsupervised Learning, Time Series Clustering, Macroeconomic Analysis, Empirical Asset Pricing, Multi-Factor Portfolio

TABLE OF CONTENTS

LIST OF FIGURES

I. Introduction

I.1 Set-up

Returns on empirically robust risk factors like size, value or momentum are time-varying[1]. Yet, positive medium-term excess returns suggest the study of particular timing systems to only harvest the 'good states'[2]. Several researchers have tackled this problem by linking factor returns to business cycles (e.g. Hodges et al., 2017), applying a variety of econometrics instruments (Miller et al., 2015) and constructing theoretically diversified multi-factor portfolios (Bender et al., 2013).

This thesis follows a somewhat different approach. Inspired by recent developments in the machine learning realm, I cluster return time series of five popular equity factors into several look-a-like sub-sequences to gain an understanding of reoccurring patterns and the related return dynamics. These patterns are then connected to macro-economic variables serving as state variables. While potentially linked to business cycle phases, this methodology gives more leeway than a pre-determined cycle fixation. Thus, it is more closely linked to the stochastically derived regime-based models of e.g. Ammann & Verhofen (2006). To the best of my knowledge, however, this is the first paper applying the concept of subsequence time series clustering to factor return time series.

Eventually, my analysis yields a relationship between macroeconomic variables and the factors' return prospects, which, in turn, allows the set-up of a dynamic factor rotation strategy. In contrast to most of the existing literature (e.g. Subbiah & Fabozzi, 2016), I solely focus on pre-filtered 'high conviction' trading signals to engage (long or short) with one or more factor-based investment strategies. For the remainder, the strategy will simply be invested in a buy-and-hold portfolio equally consisting of all factors. Thus, this approach can be seen as an addition to a passively managed equity portfolio, i.e. as a tool for tactical asset allocation (TAA)[3].

[1] See Chaieb et al. (2018), Ilmanen et al. (2019) or Grobys & Kolari (2019) for most recent (time-varying) estimates of the market, size, value, momentum and quality (profitability and investment) factor premia.
[2] The alternative approach would be to own a diversified buy-and-hold portfolio and 'smooth out the ride'.
[3] Ang et al. (2017) demonstrate that US mutual fund managers, on average, already deploy factor timing strategies successfully. Yet, with an incremental return of ca. 0.2 % p.a., the magnitude of their success is still somewhat limited.

The success of the strategy is then tested empirically out-of-sample. Therefore, I assume the position of an investment manager, who, having observed all factor returns as well as all relevant macroeconomic releases for the pre-trading period, is to act on the signals provided by the model. The performance of the clustering exercise is benchmarked vs. the static equal-weight portfolio as well as an ordered logit prediction model based on the same set of macro-economic indicators and a signal indicating out- or underperformance versus the factor's unconditional history.

Overall, this thesis touches upon three different branches of the existing academic literature: Equity return factors and their predictability, multi-factor rotation strategies as well as subsequence time series clustering. Relevant concepts are featured in chapter II in more detail. First, I present the most common rationales for the factors continued existence as well as the scientific state of the art regarding the factors' predictability, which serve as motivation for the explanatory variables selected for the empirical exercise performed later. I also introduce the concept of subsequence time series clustering and showcase its relevance for the issue at hand. Later on in chapter III, I describe the dataset and discuss the methodology used. Overall, I have six factor return time series as well as 65 monthly macroeconomic indicators spanning the period July 1993 – December 2019 at my disposal. In contrast to e.g. Amenc et al. (2003), I explicitly focus on 'academic' factors, not potentially noisy (smart beta) ETFs that are often only claiming to fully harvest the desired factor premia[4]. While this makes my strategy less investable for the retail investor, it can still easily be accessed by institutional investors via total return swaps or even structured notes. I then use the subsequence time series clustering method developed by Yeh et al. (2016) to identify repeating patterns, or motifs, in the return time series. Going through the first $^2/_3$ of the dataset - the training dataset - I identify 28 motifs, which are further validated using different testing procedures. Given the motifs, the economic regime is identified using ordered logit regressions of the transformed and one month-lagged macro-economic indicators on the phases identified by the subsequence time series clustering. Lastly, I use the most recent macro-economic indicators to determine whether the upcoming period can be exploited by the model. The yielded results are presented in chapter IV. Overall, the motif-based strategy outperforms both its simpler macro-based benchmark as well as the static equal-weight portfolio by substantial margins, gross of costs. Regressing the returns on the six factors

[4] These dedicated factor ETFs often do not fully harvest the targeted factor premia. Analyzing all US equity ETFs with at least 36 month of performance history, Blitz (2017) derives factor betas of smaller than 0.24 (targeting Momentum) and 0.88 (Size) as well as significantly negative alphas.

used to construct the portfolios, I find highly significant sensitivities to the best-performing factors and vice versa. The simpler macro model, on the other hand, is not explained well by its underlying factors. Moreover, the motif model generates significant alpha vs. the European Fama-French six-factor model, while the macro model doesn't. Incorporating conservative implementation cost estimates doesn't materially change the results. Exemplary, the increase in the Sharpe ratio is robust to five times the trading costs initially assumed. Finally, chapter V discusses the findings and leaves the interested reader with an outlook on further research.

I.2 Hypothesis

As mentioned earlier, this thesis approaches the concept of utilizing time-varying factor returns in reverse to the current academic literature. Instead of assigning parts of a factor return time series to one of several economically motivated phases, I determine return patterns free of any pre-set definition or label. In other words, I do not assume returns to behave similarly across arbitrarily set periods, but rather analyze the macroeconomic backdrop in times of similar return patterns. Thus, I test the following informal hypothesis:

Subsequence time series clustering improves the feasibility of an equity factor timing strategy in comparison with traditional macro-economic models

Several performance metrics lend themselves to this kind of analysis. Inspired by Harvey's (2017) call for more generalizable results in the field of financial economics, I use a variety of risk-return metrics. Moreover, as I do not aspire to take a position on the "correct" measure of risk, I will adhere to the two main approaches: Measures of variability of returns as well as a measure relating to an enduring sequence of losses or drawdown.

First, traditional consumption-based asset pricing models imply the use of standard deviation as a risk measure and thus the Sharpe ratio as the relevant metric. The Sharpe ratio of a given strategy *i* as a risk-adjusted performance measure is defined as follows:

(I) $$SR_i = \frac{\bar{r}_i - r_f}{\sigma_i},$$

With the mean[5] return of strategy *i* and the risk-free rate, $\bar{r}_i$ and r_f, respectively as well as the standard deviation of strategy *i*, σ_i. By setting SR_{motif} as the main strategy's Sharpe ratio

[5] I typically annualize all returns assuming 252 trading days or respectively 12 months in a calendar year.

and SR_{macro} as the traditional macro strategy's Sharpe ratio, the stated hypothesis can be re-organized to:

(II) H_0: SR_{motif} <= SR_{macro}

(III) H_1: SR_{motif} > SR_{macro}

While popular in both academic and professional circles, the Sharpe ratio only considers a fraction of the available return information. In this context, I will also make use of the Omega ratio $OmR_i(t)$ developed by Shadwick & Keating (2002), which incorporates the skewness and kurtosis of a given strategy, defined as follows:

(IV) $$OmR_i(t) = \frac{\int_t^{\infty}(1-F(x))dx}{\int_{-\infty}^{t} F(x)dx},$$

With t as the assumed threshold return, that divides the cumulative return of the return distribution, $F(x)$ is the cumulative probability distribution, $\int_r^{\infty}(1-F(x))dx$ is the probability of a return above the given threshold t, and $\int_{-\infty}^{t} F(x)dx$ is the probability of a return below the given threshold t.

Moreover, the ratios mentioned above treat upside and downside variability equally. The Sortino ratio SoR_i tackles this issue by only considering the return distribution below a minimum acceptable rate r_{MAR}:

(V) $$SoR_i = \frac{\left[(\prod_{i=1}^{n}(1+r_i))^{\frac{1}{n}}-1\right]-r_{MAR}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(\min(r_i-r_{MAR},0))^2}},$$

Second, investment professionals frequently use a strategy's maximum drawdown over a given period to judge its loss potential. More precisely, they calculate the corresponding MAR ratios to derive the risk-adjusted return of an asset. The MAR ratio is defined as the annualized cumulative return $\tilde{r_i}$ over the maximum drawdown MDD_i:

(VI) $$MAR_i = \frac{\tilde{r_i}-r_f}{MDD_i},$$

As illustrated by Martin & McCann (1989), investment decisions based on maximum drawdown instead of standard deviation can result in significantly different optimal allocations and, in turn, investment performances.

II. Literature Review

Overall, this thesis touches upon three different branches of the existing academic literature: Equity return factors and their predictability, multi-factor rotation strategies as well as subsequence time series clustering. The first part of this review thus features a concise overview of six of the most popular equity factors as well as the rationales for their continued existence. These rationales can be divided into three interlinked branches: The strategies act as a proxy for the true underlying risk exposure to risk factors like growth surprises and interest rate risk (Winkelmann et al., 2013), arise from deep-rooted behavioral or otherwise irrationally biases that are present in a meaningful fraction of investors (Barberis & Huang, 2002) or stem from institutional or market structure-related issues (Lucas et al., 2002). While the following sections will selectively feature concepts from the second and third branches for illustration purposes, this thesis mainly focuses on the first branch.

Building on this prologue, the second part of this section presents the existing empirical evidence of the factors' predictability on both an individual as well as a multi-factor basis and its implementation within a tactical asset allocation framework. It thus serves as motivation for the explanatory variables selected for the empirical exercise performed in chapter IV. Lastly, I introduce the concept of subsequence time series clustering and its applicability to the issue at hand, factor timing.

II.1 Return Factors

" There are huge dangers with computers. People calculate too much and think too little"

Charlie Munger, Berkshire Hathaway Annual Shareholder Meeting, 2002[6]

While numerous return factors have been published in academic journals around the globe[7], only a handful prove to be persistent across time, geographies and definitions. Exemplary, Beck et al. (2016) analyze the entirety of the SRRN database and only finds the low beta, value, momentum and illiquidity factors to yield positive excess returns, after controlling for transaction costs. Surprisingly, the popular quality and size factors are deemed questionable, at best. However, e.g. Asness et al. (2018a) establish a fundamentally sound link between quality and size and demonstrates that, controlling for quality, the size factor has delivered significant alphas over time. Therefore, the following exercise adapts Bender et al. (2018) by making use of all factors mentioned above. This selection is also in-line with state-of-the-art factor selection frameworks, like the one proposed by Hsu et al. (2015) and Fama &

---

[6] Quoted according to Bevelin (2005), page 52.

[7] Harvey et al. (2016) analyze a total of 296 factors established by academic researchers over the past five decades, and find between 27% and 53% insignificant given more prudently set significance thresholds. The factor universe can be found here: https://faculty.fuqua.duke.edu/~charvey/Factor-List.xlsx

French (2017) and consistent with Ahn’s et al., (2011) purely statistical component analysis, which suggests the use of six factors in a capital asset pricing model. The following section thus shortly introduces the six factors by detailing the rationales brought forward by the academic community.

Kicking off the list is the low-risk (beta/ volatility) premium, originally documented when testing the cornerstones of the CAPM, which argues that stocks with a higher beta than the equity market do not produce higher returns (Haugen & Heins, 1975). Thus, it contrasts investment theory’s most basic idea: The security market line does, in fact, not form a positively sloped curve (e.g. Black, 1972).

Although under intense scrutiny lately e.g. by Novy-Marx & Velikov (2015), Fama & French (2016) and Beveratos et al. (2014), Blitz (2016) finds that increased exposure to the market beta is not rewarded with a positive premium in the cross-section, even in the context of Fama & French's (2015) five-factor model. Moreover, Baker et al. (2011) demonstrated this equation to also hold for low-volatility stocks.

A variety of theories have been developed to shed light on the reasons for the premium’s persistence: First, investors lacking sufficient or cheap access to leverage may use high-beta stocks to increase their respective portfolio returns at the expense of a lower Sharpe ratio. Decomposing beta into correlation and volatility, Asness et al. (2018b) indeed find the resulting low-correlation factor BAC to deliver strong performance both across time and geographies. Second, asset owners regularly set institutional benchmarks, thus shifting the investor’s preference away from optimal on a standalone risk-adjusted basis towards optimal on a relative basis (Baker et al., 2011). Typically, low volatility stocks have lower betas, thus overweighting these stocks leads to higher tracking errors for institutional investors, ceteris paribus. Such tracking errors need to be justified by sufficient excess returns. Third, short-term fund flows are directed towards high-beta stocks as a result of sell-side analysts’ excessively inflated growth projections. Consequently, prices for these high-beta stocks are bid-up, expected returns diminish and the Low-Vol premium arises (Hsu et al., 2013).

Related, but distinctly different is the Value factor, i.e. the tendency of cheap stocks to outperform expensive ones over the long-run (Basu, 1977). While originally documented for strategies based on book-to-market ratios, other measures of value like the price-earnings- or the cash flow-to-price ratio have also been found to carry positive excess returns (Asness et al., 2015).

In line with Merton (1973), Fama & French (1992) demonstrate that the excess returns occur as compensation for risk, i.e. an increased sensitivity to shocks to real GDP. This is due to value stocks' balance sheet structure: Overall, these companies tend to form in capital-intense industries with more irreversible investments (Campbell & Voulteenaho, 2004) and higher financial leverage (Chen & Zhang, 1998) impeding the companies' capability to quickly adapt to unfavorable economic environments (Cochrane 1991). In periods of high corporate earnings growth, on the other hand, value stocks profit disproportionately. Furthermore, value stocks possess different cash flow characteristics than growth stocks – e.g. their cash flows tend to be shorter in duration – resulting in different loadings on changes in the slope of the yield curve (Campbell et al., 2009).

In contrast to the value factor, which can be interpreted as a trend-reversal strategy, momentum strategies deal with the exploitation of stable trends. Since at least the late 1960s, these types of trend-following are associated with significant abnormal returns (Levy, 1967), which cannot be explained by traditional asset pricing models (Jegadeesh & Titman, 1993). But why do these trends form and stay intact long enough for momentum traders to engage?

Most of the explanations laid out by the academic community are behaviourally motivated, but there is also a risk-based argument to be made for the existence of the momentum factor. Given the factors' return profile resemblance to the FX carry trade, i.e. earning a steady premium in stable markets, but crashing when the 'tides turn', Daniel & Moskowitz (2016) develop an option-based framework and demonstrate momentum's long convexity structure: In general, a stock can be interpreted as a call option on the company's assets when the company carries debt on its balance sheet (Merton, 1974). Given periods of market stress, the option of a company that has suffered most severely – i.e. the short leg of the market-neutral momentum trade – features a significantly increased gamma. On the flip side, the gamma of the long leg is still suppressed, rendering the strategy long gamma and thus exposed to sudden and intense reversals.

Moving on, the size premium illustrates the promising excess returns of small-cap over large-cap stocks. Beginning with the seminal contributions by Banz (1981) and Reinganum (1981), which regard the size premium as a compensation for the gathering and interpreting information, the community has developed plenty of additional hypotheses on the premium's existence.

Exemplary for a plethora of subsequent research, Fama & French (1995) relate relative financial distress and other variables linked to changes of the available investment opportunity set (Petkova, 2006) or of the economic environment (Vassalou and Xing, 2004) to the performance of the size effect over time. The latter argument was developed further by Campbell et al. (2008), who demonstrate that US firms with a high probability of bankruptcy have a high loading on the size factor. More generally, Winkelmann et al. (2013) find size portfolios react most sensitive to real economy shocks in their VAR models. Other risk-based explanations include the 'fallen angels' phenomena, which argues that small firms generally have lost market value due to their bad performance (Chan & Chen, 1991).

Hou & Moskowitz (2005) tackle the size premium from a market structure perspective and find delayed price responses to account for a large part of the size. With the gradual digitalization and sophistication of financial market participants, some papers thus argue that size is no longer relevant post-1980, yet Asness et al. (2018a) demonstrate that when controlled for the quality of a firm, the 'clean' size premium is similar in economic significance to value and momentum[8].

In this spirit, the quality factor aims to capture the excess returns of high-quality companies over companies of lower quality. Although this has been a widely adopted concept in fundamental analysis since Benjamin Graham's book on Security Analysis (1934), it's a relatively new phenomenon in quantitative investments. The main challenge to date is to consistently define the quality factor using quantitative indicators. Commonly a firm's quality is associated with its competitiveness, efficiency and return-on-equity (Dechow et al., 2010) as well as its accounting quality, pay-out and dilution (Hsu et al., 2019). Fama & French (2015) proxy quality using two of these metrics - profitability and investment – in their seminal five-factor asset pricing model. Moreover, the quality factor as estimated by AQR via their 'Quality Minus Junk' indicator (Asness et al., 2018c) explains a significant part of the variation of the investment returns of one of Graham's most prominent scholars – Warren Buffet (Frazzini et al., 2018a).

As there is no standard definition of the quality factor, no unified theory for its continued popularity exists. There is, however, an immense amount of research covering some of the indicators mentioned above[9]. First, the profitability premium - more-profitable firms earn an excess return versus less-profitable firms (Ball et al., 2015) – can be explained from a

---

[8] Even more recently, Ciliberti et al. (2019) find the size factor to be highly significant (t-stat of 5) when controlling for exposure to the market and low-risk factor.

9 Commonly used characteristics, like low book leverage or low earnings growth volatility are more related to the low-risk than the quality factor (Hsu et al., 2019).

mispricing perspective: Investors tend to underreact to high profitability ratios because of biased beliefs based on salient information like EPS growth or momentum (Bouchaud et al., 2016) as well as the profitability measure's complexity relative to other ratios. The latter claim is further substantiated by the observation that portfolios constructed on less-manipulated proxies for corporate profitability exhibit lower returns than the ones of more complex proxies like Novy-Marx's (2013) gross profits-to-book assets ratio.

Related to measures of profitability is a firm's accounting quality, i.e. how aggressively sales, that have not yet (or may never) translate into actual cash flows, are being accounted for. Firms boosting current earnings by accumulating high accruals were shown to feature relatively low subsequent returns (Chan et al., 2006). Hirshleifer et al. (2004) attributed this effect to investors' focus on headline earnings as well as ignorance of possible manipulation of those earnings.

Moreover, corporate management has some discretion on how to allocate the firm's net earnings. Which share of profits is paid out via dividends as well as, more importantly recently, buybacks or negative buybacks (i.e. share issuances) and how should the remains be re-invested in the business? The empirical results are straightforward: While firms that pay out more have higher subsequent returns (Boudoukh et al., 2007), most forms of share issuance lead to underperformance (Loughran & Ritter, 1995). In terms of investment, Cooper et al. (2008) demonstrate the appeal of modesty: Neither too little nor too much investment results in superior stock returns. There are two explanations for this. First, firms that can finance a high level of investment must be deploying capital into safer projects, which tend to produce lower returns (Hou et al., 2015). Secondly, firms might overinvest out of the misalignment of managerial incentives and the consequential 'empire building' (Roll, 1986).

Lastly, the expected cost of trading assets across all market phases, or, to put differently, an asset's liquidity, raises the expected return investors require from holding it, thus carrying a risk premium (Amihud & Mendelson, 1986). As idiosyncratic liquidity can be diversified away in a portfolio context, the liquidity factor is thus a function of an asset's exposure to systematic liquidity shocks (Pástor & Stambaugh, 2003). These shocks arise out of (sudden) changes to both supply and demand: First, the availability of funding to liquidity providers like dealer-brokers matters (Kondor & Vayanos, 2019): When funding conditions are tight, it is more costly for traders to take on new positions, in particular capital intensive positions in high-margin securities. This decreases overall market liquidity and may lead, in the worst case, to 'liquidity spirals' in which traders react to falling asset prices by liquidating additional positions, amplifying the spiral in the process (Brunnermeier & Pedersen, 2009). Second, demand-side considerations like one-sided and concentrated trading behavior of institutional investors (Koch et al., 2016), high pairwise correlations of individual securities and thus weak incentives to trade in individual securities (Morck et al., 2000), as well as changes to investor sentiment in the context of 'panic selling' (Hameed et al., 2010) were shown to substantially affect market liquidity.

Given the time-varying nature of the mentioned indicators of liquidity, the magnitude of the liquidity premium can vary significantly over time as well. While there are a couple of different approaches to estimating the liquidity factor, Pástor & Stambaugh's (2003) recently successfully replicated10 method is still one of the most popular ones11: Building on Amihud's (2002) liquidity measure of the average daily ratio of absolute stock return to dollar volume, Pástor & Stambaugh suggest the equal-weight average of all stocks' Amihud liquidity to equal the market liquidity. In turn, they create the liquidity factor by creating a portfolio that is long the decile stocks most sensitive to market liquidity shocks and short the decile stocks least sensitive. This factor is often associated with the size and momentum factors but was shown to also appear in the cross-section of returns when controlling for different measures of the two factors (Asness et al., 2018a).

In summary, my literature research indicates that the low-risk, value, momentum, size, quality and liquidity factors are all strongly supported by academic research and can be linked to macro-economic risk factors. The following section thus delves deeper into the factors' time-series predictability and gives an overview of potential predictor variables used in section III.2.2.

II.2 Predictability and timing models

Given its relevance for investment decisions, return prediction is a popular topic in both academic and industry circles. Beginning in the late 1980s, researchers like Keim & Stambaugh (1986), Fama & French (1989) and Hodrick (1992) discuss models to predict long-term aggregate stock market returns. Commonly used indicators were the dividend yield, short-term interest rates, the term spread, and credit spreads. Later on, several authors extended the range

[10] See Pástor & Stambaugh (2019) for a discussion of the two replication studies by Li et al. (2019) and Pontiff & Singla (2019).

11 Most recently, Ciliberti et al. (2019) calculate a Cold-Minus-Hot indicator based on the average daily dollar volume of transactions and finds it to be highly significant, yet essentially equivalent to Pástor & Stambaugh's indicator controlling for their different low-vol exposure.

of valuation metrics by estimating price-to-book and price-to-earnings ratios and by adjusting these ratios to fluctuations in the business cycle (e.g. the Shiller CAPE ratio)[12]. While a blend of metrics shows the highest predictive power, the Shiller CAPE ratio is viewed as the best single predictor (e.g. Keimling, 2016).

Research on single factor prediction models, on the other hand, is still scarce, with only a handful of papers published so far. This line of research can be divided into three distinct branches, focused on valuations[13], sentiment or macroeconomic indicators.

Beginning with valuations, Arnott et al. (2016) and Asness et al. (2017a, 2017b) discuss an important point: Just like with any asset, the attractiveness of a factor-based strategy should also be measured based on its current valuation. Using the price-to-book ratio and size, value and momentum portfolio returns for the period January 1967 to September 2015, they derive significantly negative correlations between the valuation measure and each factor's subsequent 5-year return.

Given that the valuation of a factor can be regarded as a function of its past momentum – high valuations typically follow previous price increases - momentum-based models were developed. Exemplary for a large swath of academic papers, Bird & Casavecchia (2007) or Leivo & Pätäri (2011) illustrate the positive effects of incorporating price and earnings trend data on value and growth strategies' Sharpe ratios. More recent approaches include multi-factor portfolios constructed according to the individual factor momentum (Gupta & Kelly, 2019) or the momentum of the optimal factor weights in the portfolio (Leippold & Rüegg, 2019).

While there is some evidence14, that investor sentiment around a particular stock is significantly impacted by the stock's momentum, sentiment can also be regarded as an investor's inclination to trade. In this context, e.g. Baker & Wurgler (2006) construct a composite sentiment index out of six underlying proxies for sentiment[15] and find the returns of size, high-risk and low-quality portfolios to increase after periods of low sentiment. Specifically, when sentiment is low – i.e. the price for risk is high - subsequent returns are

---

[12] Campbell & Shiller's (2001) model on forecasting long-term equity market returns remains a seminal model to this day.

[13] The importance of valuation also features a systematic component: As illustrated by Rothman (2007), markets exhibiting crowding in similar factors can experience phases in which factor returns no longer reflect the underlying macroeconomic risks, but rather forces of market structure-based supply and demand. The resulting vicious circle can lead to increased systematic risk and LTCM-like scenarios.

[14] See Uhl et al. (2015) for a discussion of the underlying chicken-hen theme: Does sentiment influence price momentum or vice versa?

[15] Find the monthly time series here: http://people.stern.nyu.edu/jwurgler/

relatively higher for unprofitable, distressed, non-dividend-paying, young, small, highly volatile and extreme growth stocks. Moreover, measures of risk aversion like the VIX Index for implied aggregate volatility were shown to have some forecasting ability for the size, value and quality factor (Copeland & Copeland, 1999) as well as the momentum factor in particular (Daniel & Moskowitz, 2016).

Lastly, several papers deal with the predictive power of macroeconomic indicators like industrial production or inflation rates because of these indicators' capability of describing the current state of the business cycle, i.e. the magnitude of non-diversifiable, systematic risk (Jacobs & Levy, 1996). Other commonly used indicators include both short- and long-term (and their spread) interest rates on government bonds (Conrad et al., 2003), the spread differential between government and corporate bonds (Chan et al., 1985) and exchange rates (Cenedese et al., 2016). Moreover, Conrad et al. (2003) use trends in corporate earnings measured via past stock market returns to estimate future returns of size and value portfolios. Analyzing a large variety of potential predictor variables, Parker et al. (2016) find trailing 1-month S&P 500 returns, trailing 3-month US Dollar percent changes, trailing 12-month relative performances of the cohort (momentum of factor) and dispersions of stock returns within the cohort most frequently helpful in timing individual factor exposure. Yet, the expected utility gain varies across different investor types (Levis & Tessaromatis, 2004).

With Levis & Liodakis (1999) exploring the potential of style rotation strategies assuming perfect foresight and finding a hit ratio of at least 65% (80%) for size (value) portfolios and Ahmed et al. (2002) demonstrating that multi-style rotation can add more value than single-style rotation, research graduated from single to multi-factor models. Exemplary, Amenc et al. (2003) select the most efficient combination out of a vast list of macro-economic indicators via the models' Schwartz information criterion and implement a profitable trading strategy via style ETFs. Other statistical tools to overcome the time-varying nature of factor premia include support vector regressions (Nalbantov et al., 2006), regime-switching Markov Chain Monte Carlo models (Ammann & Verhofen, 2006) or nonparametric forecasting models (Subbiah & Fabozzi, 2016). Most recently, Campbell et al. (2009) and Lochstoer & Tetlock (2020) use panel VAR techniques to forecast firm-level expected returns, then aggregate the estimates into portfolios. Haddad et al. (2020) extend this approach by explicitly imposing structures on the implied pricing kernel and find a variety of equity factors to be strongly and robustly predictable. They do not, however, apply their framework to a dynamic factor switching trading model.

In summary, a large body of research from both academic and industry circles suggest the profitable use of return prediction models in the context of a dynamic factor rotation strategy. Yet, as mentioned in the introduction, my method approaches the concept of utilizing time-varying factor returns in reverse to the current academic literature. Put differently, I analyze the macroeconomic backdrop in times of similar return patterns, not the other way around. These patterns are identified using a technique called subsequence time series clustering, which I introduce in greater detail in the following section.

II.3 Subsequence Time Series Clustering

The academic literature has developed a variety of models and tools to identify distinct patterns both within a single and across multiple time series. This section focuses on the progress made in the field of unsupervised learning of a hidden data concept, or more precisely, cluster analysis. In short, cluster analysis deals with the identification and subsequent grouping of similar data objects without requiring advanced knowledge of the particular group's definition. Thus, unlike explicit data mining tools like e.g. Random Forests, cluster analysis does not aim at forecasting a target variable (Rai & Singh, 2010)[16].

While originally developed to group a data set according to static criteria (Driver & Kroeber, 1932), clustering techniques have also been adopted to the more-dimensional space like time series analysis (Aghabozorgi et al., 2015). The latter strand of the literature can be roughly classified into three categories:

1) Whole time-series clustering
2) Subsequence clustering
3) Time-point clustering

In contrast to the first concept, which aims at clustering multiple time series according to their similarity, the latter two are concerned with only one time series at a time (Zolhavarieh et al., 2014). Moreover, as I am interested in the particular return dynamics of every factor by itself over a given period, this thesis is focussed on subsequence time series clustering (STSC).

Conceptually, STSC groups similar 'segments' of a given time-series according to their similarity.

[16] Given this paper's focus on timing factor exposure, there is, naturally, a forecasting component to this work. Please refer to the economic regime identification in III.2.2 for more details.

The following figure schematically illustrates the idea:

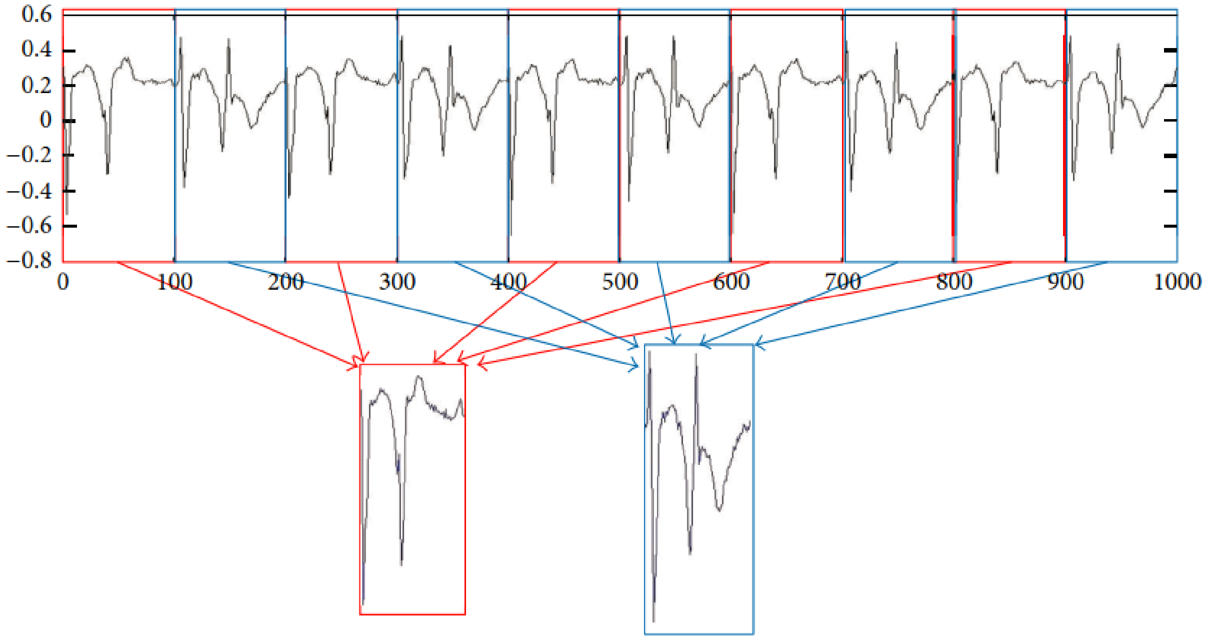

Figure 1: A sample of subsequence time series clustering, taken from Zolhavarieh et al. (2014), p. 3

Given the time series T depicted in the upper panel over the $0 < t < 1000$ period, the algorithm selects ten shorter sub-sequences $T_{i,k}$ with the length $k = 100$ starting at position i and identifies two distinct patterns. In the second step, the respective subsequences are fully assigned to either the red or blue pattern. Thus, a hard partitioning clustering algorithm like the popular k-means method is used. While this approach gained popularity in the 1990s and early 2000s, it was shown to only produce sine wave-like results independent of the input time series (Keogh & Lin, 2005). As indicated by Rodpongpun et al. (2012), this problem can be solved, if some data is considered noise and afterward blatantly ignored. The following figure illustrates the approach:

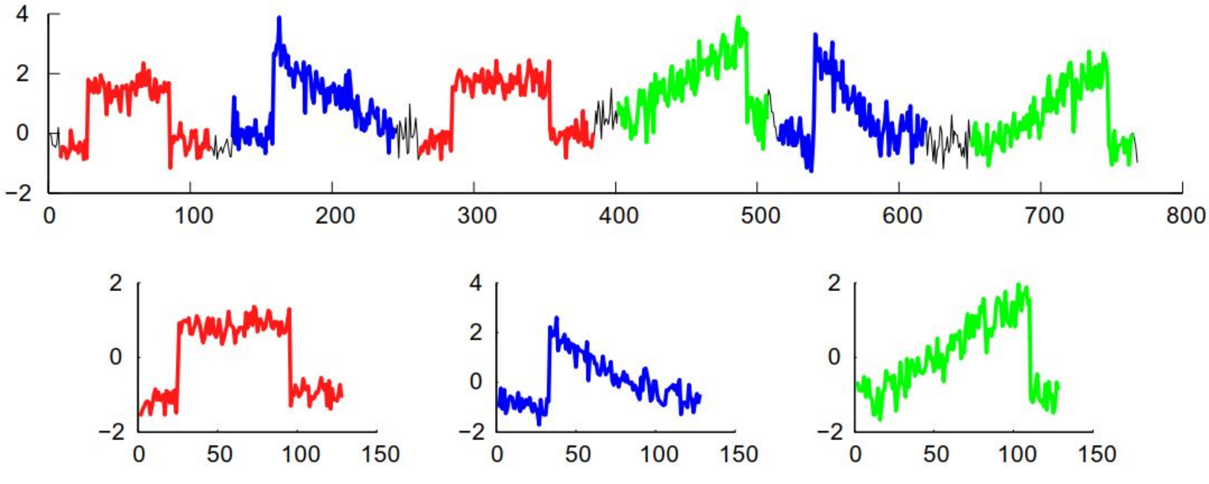

Figure 2: Illustration of selective subsequencing, taken from Rodpongpun et al. (2012), p. 6

While the algorithm shown in Figure 1 aims at fully clustering the entire sequence, the algorithm used to plot the figure above only considers selected subsets of the time series – called motifs - informative and, in turn, generates three meaningful cluster centers colored red, blue and green. These cluster centers represent their respective cluster members so that important features and shapes are preserved. This approach is particularly helpful when identifying motifs within a very noisy time series like return time series. Rodpongpun's et al. (2012) approach exhibits three features:

1) Motif discovery to extract subsequences & create clusters
2) Subsequence matching to add subsequences to existing clusters
3) Brute force to merge two existing clusters into one

Yeh et al. (2016) demonstrate that the calculation of these three steps can be significantly accelerated using a meta-time series called the matrix profile. Essentially, the matrix profile is a vector consisting of the z-normalized Euclidean distance (ED) between each subsequence and its corresponding nearest-neighbor. It is derived as follows: Given a time series $T = t_1, t_2, \dots, t_n$ where n is the length of T, and each subsequence $T_{i,m} = t_i, t_{i+1}, \dots, t_{i+m-1}$ starting at position $1 \leq i \leq n - m + 1$ is of length $m < T$, there are $n - m + 1$ subsequences – overlapping subsequences are valid - of length m.

To find each subsequence's nearest neighbor, the z-normalized ED, originally developed by Faloutsos et al., (1994), is calculated as follows: Given two subsequences $t_i = t_{i,1}, t_{i,2}, \dots, t_{i,m}$ and $t_j = t_{j,1}, t_{j,2}, \dots, t_{j,m}$ also with $1 \leq j \leq n - m + 1$, their z-normalized transformations are $\widehat{t_{i,a}} = \frac{t_{i,a} - \mu_{t_i}}{\sigma_{t_i}}$ and $\widehat{t_{j,a}} = \frac{t_{j,a} - \mu_{t_j}}{\sigma_{t_j}}$, where $1 \leq a \leq m$. In turn, the z-normalized Euclidean distance $d(t_i, t_j)$ between t_i and t_j is defined as follows:

$$\text{(VII)} \quad d(t_i, t_j) = \sqrt{\sum_{a=1}^{n} (\widehat{t_{i,a}} - \widehat{t_{j,a}})^2}$$

Visually, the Euclidean distance can be thought of as the distance between the observations of two time series t_i, t_j at any point $1 \leq a \leq 50$ in time, which, in the following example amounts to 37.18. While the absolute value has no further interpretation on its own, it can be compared to EDs at other points of time.

The following figure visualizes this example:

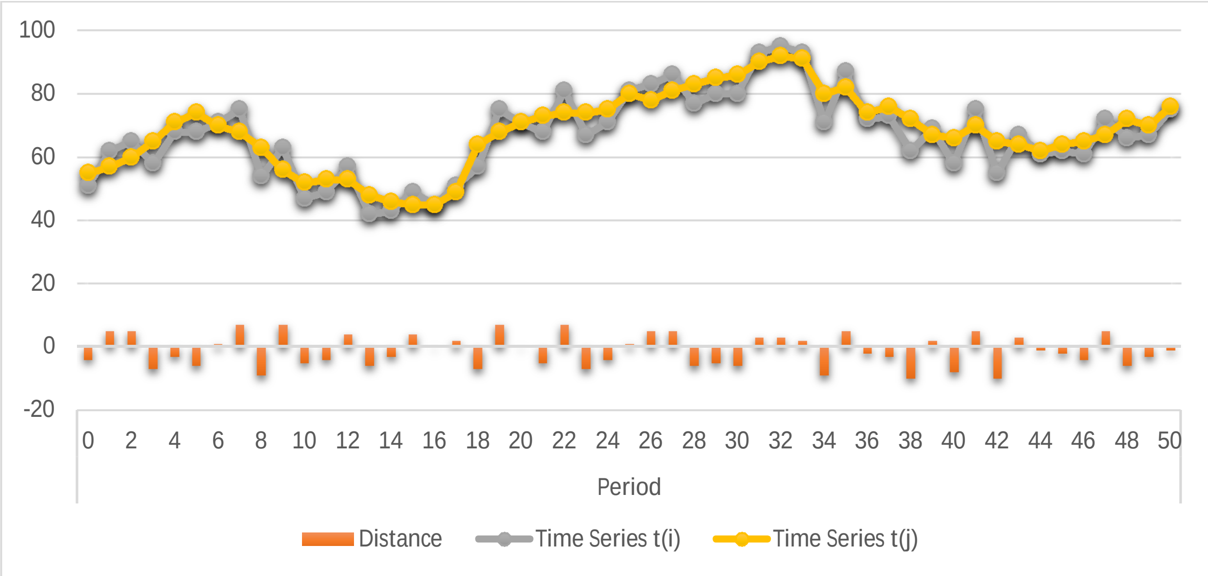

Figure 3: Illustration of Euclidean distance measure

All pairwise Euclidean distances can be stored in a new matrix M, which spans $(n - m + 1)\ x\ (n - m + 1)$ in dimension. To reduce the matrix's massive dimensionality without losing crucial information, the matrix profile is calculated via randomly iterating m-times through the matrix M (without repetition) and selecting the minimum figure elementwise, starting with a positive infinity vector (inf). This process is shown in the following figure, which depicts the current value of the matrix profile (top box) with the third (bottom box, left figure) and twelfth row (bottom box, right figure) of a representation of matrix M. Note that values of '0' (highlighted yellow) are skipped.

Iteration step I. Iteration step II:

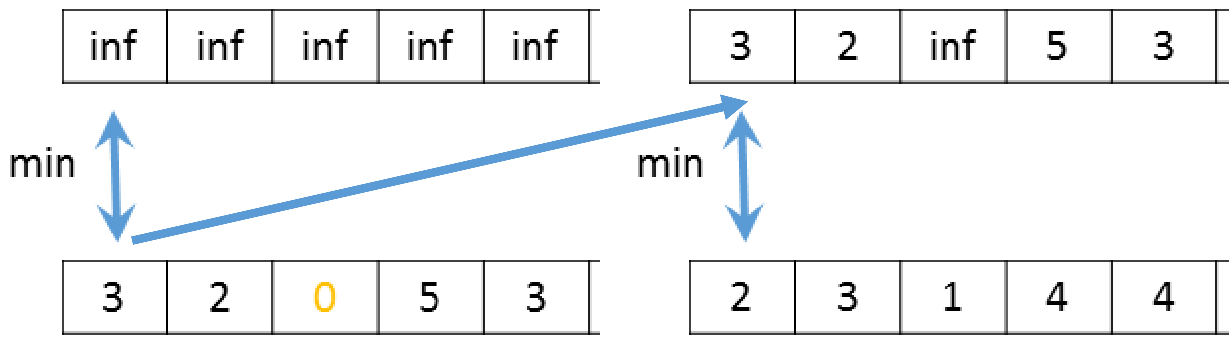

Figure 4: Illustration of applying pairwise minimum to a representation of matrix M

After successfully iterating through all rows, the matrix profile at location i is the ED between T_i and its nearest neighbor. Thus, as lower matrix profile values imply closer resemblance to other subsequences, local minima (e.g. observation 4.200) of the matrix profile shown below can be interpreted as motifs, and local maxima as discords (e.g. observation 3.800):

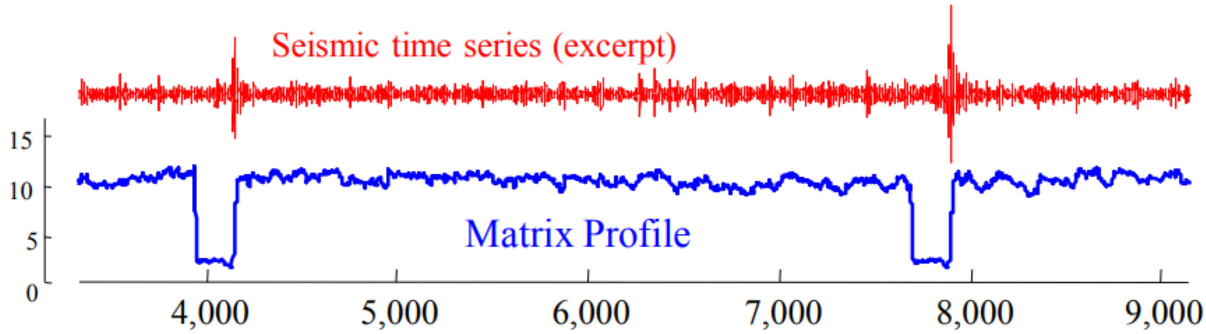

Figure 5: Illustration of the matrix profile representation, taken from Yeh et al. (2016), p. 6

As elaborated on in more detail in section III.2, I will make use of both pieces of information, i.e. motifs and discords in the application of this approach to my factor return time series.

Summing up, subsequence time series clustering can be a very useful tool when looking to identify patterns or motifs in a dataset. Using the proposed matrix profile algorithm of Yeh et al. (2016), the otherwise processing-intense computation can be reduced significantly, while the most important information is preserved.

III. Data and methodology

III.1 Data

The following chapter presents the available dataset of factor returns and macro-economic indicators and details the applied transformation methodology.

For the first, I gather the return data for the European low-risk, value, momentum, size and quality (QMJ) factors from the website of AQR[17]. As AQR also calculates a value factor in the spirit of Fama & French (1995), I am left with a dataset comprised of six factors spanning the July 1993 – December 2019 period18. Therefore, the dataset covers both periods of strong economic expansion and contraction.

However, as discussed by Bender et al. (2018), the systematic exposure of an equity factor – of momentum in particular – is prone to material changes over time. As this paper aims at developing a tool for non-directional/ market-neutral tactical asset allocation - not for strategic or directional/ delta-one asset allocation - this feature demands some pre-clustering cleaning.

To avoid gaining market exposure through the metaphorical backdoor, I beta-neutralize all return time series by adopting the recommendations set out by Cenesizoglu et al. (2016): Estimate rolling 12-month beta to daily time series and rolling 5-year beta to monthly time series as the best guess for the factor's ex-ante beta. To achieve homogenous time series length, though, I also use 12m-betas for monthly time series for the first five years. Furthermore, I beta-adjust the return time series for the first year according to their realized beta over that time frame[19].

While this adjustment process i) does not fully eliminate the factor's market beta and ii) adds some (estimation) noise to the model, it a) is more exact than simply adjusting the factor

---

[17] There are also at least four universities (CFR Cologne, Dartmouth College, UZH Zurich and SNS Pisa) as well as a variety of commercial providers (e.g. MSCI, FTSE and Eurofidai) that calculate European factor indices. These indices do, however, neither cover the entire period or investment universe nor are available free of charge.

[18] Because there is no publicly available data on the European liquidity factor, this factor is omitted from the analysis from this point onward. In general, publicly available liquidity data is fairly scarce: I am only aware of the US-based liquidity factor data available on Prof Pastor's website.

[19] I am fully aware that this adjustment might give the impression of me accidentally creating a look-ahead-bias. Yet, as the trading strategy does not trade in the first year, I can think of no reason, why I shouldn't adjust the time series by its 'true' realized beta.

by its long-term beta and b) does not alter the strategy's general execution feasibility[20]. I discuss the last point in more depth in section III.2.3. Essentially, beta-adjusting the factors allows the frictionless implementation of this framework into a market-neutral sub-portfolio designed to serve as an add-on to a strategic delta-one allocation.

The beta-adjusted factors generate the following individual risk-return profiles, without taking trading and rebalancing costs into account:

	Low-risk	*Value FF*	*Value AQR*	*Momentum (UMD)*	*Size*	*Quality (QMJ)*
Mean Return p.a.	12,9%	2,6%	0,3%	13,1%	1,5%	7,4%
Compound Annual Rate of Return (CARR)	13,0%	2,6%	0,1%	13,1%	1,4%	7,5%
Annualized Volatility	8,8%	5,2%	6,1%	8,9%	6,0%	4,9%
Ann. Downside Volatility	5,2%	3,1%	3,6%	5,8%	3,8%	2,9%
Skewness	-0,2	0,2	0,7	-0,7	-0,4	-0,1
Excess Kurtosis	2,7	6,4	9,9	10,4	3,0	4,7
Maximum Drawdown	-50,2%	-32,5%	-48,3%	-36,3%	-43,0%	-16,4%
Sharpe ratio	1,5	0,5	0,1	1,5	0,3	1,5
Sortino ratio	2,5	0,8	0,0	2,2	0,4	2,6
Omega ratio	1,3	1,1	1,0	1,3	1,0	1,3
MAR ratio	0,3	0,1	0,0	0,4	0,0	0,5

Figure 6: Descriptive statistics for European equity factor return estimates by AQR, based on daily data from July 1993 to December 2019

There are some observations worth highlighting: First, even adjusting for the factors' market beta, each factor has generated a positive excess return over the entire sample period. This result also speaks to the qualitative literature review performed in section II.1. Yet, the more traditional factors like size and value materially underperform the more recent ones like quality, momentum or low-risk on all measures. As interest rates have fairly continuously declined over the past 30 years, this performance could be due to a sample bias inherent in the dataset though. Alternatively, this result could be a function of the post-publication decay, which McLean & Pontiff (2016) estimate to amount to ca. 58% of academically acclaimed anomalies' excess returns. Second, the two different estimations of the value factor show

[20] Obviously, the beta-adjusted factors are still not 'pure', i.e. free from exposure to other risk factors like value or even country and industry affiliations. However, adjusting accordingly would make the portfolio construction way less intuitive and the trading strategy a lot less feasible, so I leave this to future (bottom-up motivated) research.

considerable differences in return, but very similar measures of risk. This is somewhat surprising given the meaningful differences in methodology[21]. Third, the momentum factor shows the most negative skewness of all factors analyzed, indicating its long-convexity profile. The factor value AQR, on the other hand, demonstrates the highest skewness. This contrast hints at the trend-following vs. trend-reversal discussion in section II.1. Lastly, the substantial positive excess kurtosis of the factors' returns point to a non-standard distribution. This stresses the need for risk-adjusted return measures that can incorporate higher moments like the Omega ratio. Exemplary, ranking the factors by their Sharpe or MAR ratio yields different results than sorting by their Omega ratio. I am left with a dataset comprised of six factors spanning the July 1993 – December 2019 period.

The macro-economic data is collected from Barclays, Bloomberg, Datastream, the FRED Economic Database as well as Global Financial Data (GFD). Inspired by the literature review performed above in chapter II.2, I obtain monthly time series for the following parameter: Industrial production, consumer confidence, inflation rates, real M1 money supply, several currency pairs like USD-EUR, USD-JPY or USD-GBP as well as broad REER-equivalents for the US-Dollar and Euro, short- and long-term government bond yields, interbank rates, yields of AAA, BBB as well as high yield corporate bonds rated lower than BBB-/ Baa3, the VIX Index and the 30d realized volatility of the S&P 500 Index. Using these inputs, I further calculate the Ted spread, i.e. the spread between the 3m interbank rate and the 3m government bond and the term premium, i.e. the spread between the yield of the 2y and 10y government bonds. If the European aggregate is not available for the entire 1993-2019 period, I estimate a linear regression between the time series of the mean of Italy, France, Germany, Spain and the UK vs. the available EMU time series and use the resulting best-fit equation parameters as a proxy for the European aggregate. Moreover, I calculate the Copeland & Copeland (1999) proxy for the price of risk, titled $risk_t$, as follows:

$$\text{(VIII)} \quad risk_t = \frac{VIX_t}{\sigma_{SPX,t,t-30}},$$

With VIX_t as the value of the VIX Index at time t and $\sigma_{SPX,t,t-30}$ as the annualized volatility of the S&P 500 Index over the preceding 30 trading days.

[21] AQR modified two features of Fama & French's methodology: First, the factor value AQR is based on current total market value of equity at the end of each month, but the factor value FF is based on total market value of equity at fiscal year-end. Second, the size and book-to-market breakpoints for the former are refreshed every month, while the latter only updates those values in June of each calendar year (Frazzini & Pedersen, 2014).

I also use stock market valuation data from the websites of Prof Goyal and Prof Shiller. Although their datasets only show metrics of US-based companies, Rapach et al. (2013) demonstrate it can still be useful in predicting developed non-US equity returns.

Because models using business cycle indicators exhibit the overall best prediction quality (Lucas et al., 2002), I also adopt several business cycle indicators from ECB (2011): As business cycles across developed countries share several common characteristics (Kose et al., 2003), I make use of a global (US, EU, Japan) set of indicators like the unemployment rate (Elsby et al., 2016), prices for commodities like gold, copper, oil and the IMF's global commodity price index, a broad commodity index (Chauvet & Potter, 2000) or economic sentiment indicators like the Manufacturing Report on Business by the Institute of Supply Management (Cecchetti et al., 1997) and the Ifo business expectations indicator for Germany. Given that France serves as a better EU-wide anchor point than Germany (Ahlborn & Wortmann, 2018), I also incorporate the INSEE business expectations for France time series from INSEE's website.

Further, as several papers have stressed the importance of the level of credit available to households and non-financial corporations (Koopman & et al., 2017), I also use the amount of bank credit of US commercial banks from the FRED database. To the best of my knowledge, this is the only cross-regional credit indicator available monthly. Lastly, López-Salido et al. (2017) demonstrate the predictive power of the share of high-yield yield bonds to total bond issuances. As there is no dedicated index showing the share of regional or even aggregate high-yield bond issuances, I created one myself using Bloomberg's SRCH function. Searching for all active and matured non-financial corporate bonds issued between December 31st, 1989 and January 1st, 2020, I found 329.970 securities with a total amount outstanding of ca. USD 19.5 trillion. Repeating this exercise, but narrowing the universe of companies to the ones with a (historic) credit rating of less than BBB- by one of the three rating agencies Fitch, Moody's and Standard & Poor's[22], I got 59.063 securities with a total amount outstanding of USD 3.5 trillion. Following López-Salido et al. (2017), I performed this exercise for global as well as US aggregates and subsequently calculated the respective ratios of high-yield bonds to total bond issuances month-by-month.

[22] Note, López-Salido et al. (2017) use data from the Mergent Fixed Income Security Database and only include bonds rated high-yield by Moody's. Therefore, my approach focusses on a broader set of bonds from a potentially smaller investment universe.

In total, I use 65 indicators for describing phases of similar factor returns, which, considering the 1.200 macro variables used by Subbiah & Fabozzi (2016), doesn’t seem too excessive.

III.2 Methodology

The following chapter details the methodology used to i) find motifs within the individual factor return time series, ii) link these motifs to the set of macro-economic indicators discussed in section III.1 and iii) set-up the dynamic rotation strategy.

III.2.1 Subsequence Time Series Clustering

Methodologically, I identify reoccurring motifs in my return time series using a Matlab implementation of the matrix profile algorithm. The algorithm requires three input vectors:

1) The beta-adjusted return time series $T = t_1, t_2, \dots, t_n$ where n is the length of T: To properly account for possible trends and autocorrelation, the time-series is also z-normalized using a 12-month window as the timeframe for calculating both the mean and the standard deviation.
2) Subsequence length m: As suggested by Bender et al. (2018), I set $m = 12$ months.
3) The maximum Euclidean distance ED_{max} that constitutes a stable pair was set at $ED_{max} = 0{,}75$. That is in-line with a lot of other studies, also from wildly different fields (e.g. Forina et al., 2009).

As introduced in section II.3, the matrix profile algorithm proposed by Yeh et al. (2016), represents an efficient and accurate method for discovering motifs in a noisy time series. While some improvements of the clustering method are currently being discussed in the academic literature, most of them, like the k-d tree algorithm, simply trade space with time (Tan et al., 2017). Moreover, recent empirical evidence suggests that the simple nearest-neighbor classification algorithm is exceptionally difficult to beat (Wang et al., 2012). As my dataset is already well-reduced in dimensionality, I stick with the proposed nearest-neighbor algorithm.

In this exercise, I, therefore, transform similar and stable pairs into consistent motifs by allowing pairs to extend over a longer period, i.e. when pairs overlap in both periods, I cluster them into one common motif. Note, however, that as of the matrix profile’s reduced complexity, I am not able to cluster the pairs according to their relative ED similarity, but rather according to their time overlap with more promising candidate pairs. Exemplary, as the two stable pairs $[t_2, t_{150}]$ and $[t_9, t_{155}]$ overlap, they are clustered into one motif m_1 spanning $t =$

[2,3, … ,21,150,151, … ,167]. The following illustrates the results of the training dataset using AQR's momentum factor UMD as an example. Note, as I exclude duplicates, pairs only show up above the blue straight line shown below.

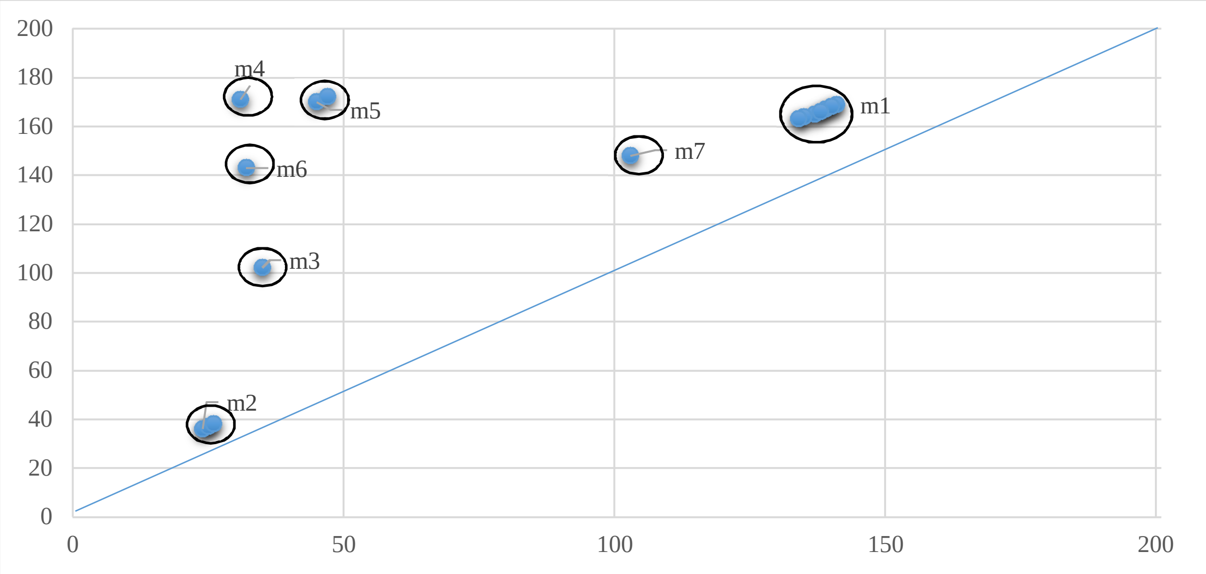

Figure 7: Illustration of the nearest-neighbor algorithm

In total, seven stable motifs (encircled) can be identified, with the motif group m_1 = [134,135, … 141,163,164, … ,169] representing the longest and m_6 = [32,143] representing one of the shorter motifs.

The more important battleground, however, is in the distance measure (Batista et al., 2013). As mentioned above, Yeh et al. (2016) utilize the z-normalized Euclidean distances (ED) as their workhorse. This comes with a big disadvantage versus other elastic distance measures like dynamic time warping, DTW (Berndt & Clifford, 1994): In contrast to ED, DTW can deal with distortions in time, i.e. lags or different speeds. This feature is highlighted in the following figure:

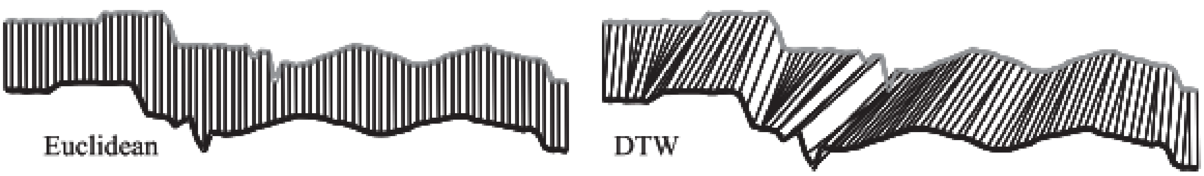

Figure 8: Differences between the Euclidean and the dynamic time-warping distance measure, taken from Keogh & Ratanamahatana (2005), p. 2

While ED measures the distance on a one-to-one basis, i.e. the distance between two points at a given point in time, DTW estimates the shortest distance between time series at a given point in time and a variety of points in time.

Many authors validate the efficiency and accuracy of DTW in conjunction with the nearest-neighbor algorithm for a variety of datasets from very different fields, like e.g. climatology, physiology or botany (Bagnall et al., 2017). Yet, there is currently no published version of the matrix profile framework including DTW or any other elastic distance measure available. I did perform two validity checks of my results, though:

First, I cross-checked the identified motifs using the Matlab implementation of dynamic time warping and found no major differences between the stable pairs identified using ED and DTW. Second, I acquired a non-published version of the DTW-matrix profile algorithm directly from Prof Keogh and compared the results. Again, I found no substantial differences to the ED-matrix profile motifs. Note, however, that because of the processing-intense computation of the DTW measure, the algorithm only yields the top pair.

Next, I explore the clustering results quantitatively by comparing their Sharpe ratios to all available periods of length $m = 12$. Exemplary, the following figure shows the distribution Sharpe ratios of all periods of length $m = 12$ and the seven identified motifs for AQR's Momentum factor UMD. Note, the cumulative distribution function (orange line) is rotated at its median $\tilde{p} = 1{,}5$ to allow for easier visualization of the distribution's descriptive statistics like median, skewness and kurtosis.

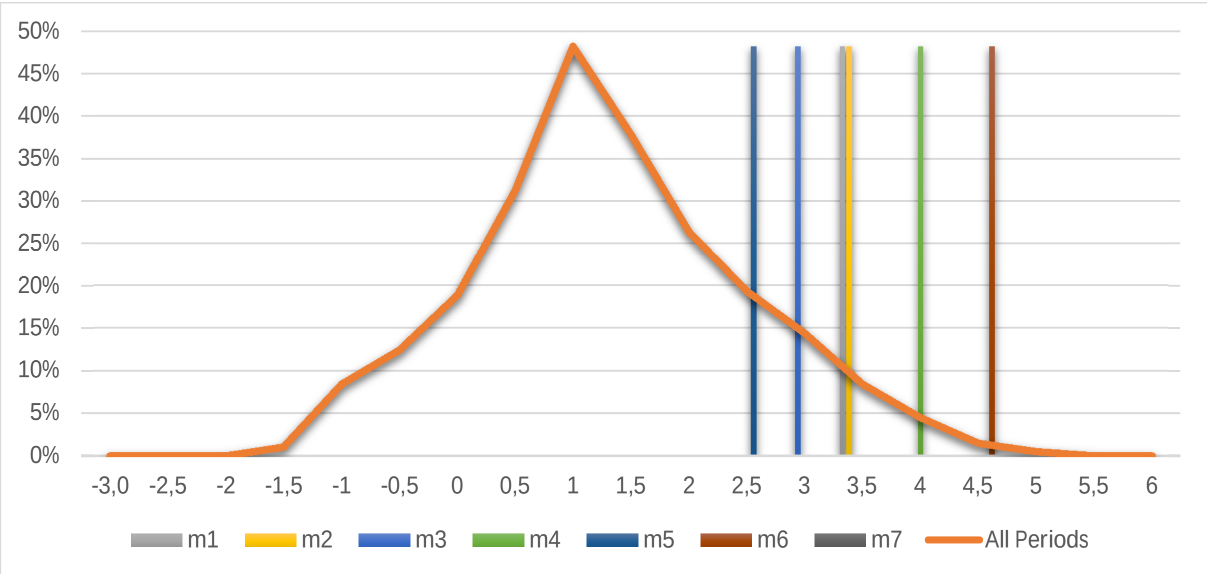

Figure 9: Validation check of identified motifs in AQR's Momentum factor using distribution analysis

There are two observations worth highlighting: First, over the training period 1993 to 2011, the European momentum factor rewarded investors handsomely in terms of (before-fees) risk-adjusted returns. Investors suffered negative 12-month Sharpe ratios in only 12.3% of all 201 observations but could have expected a median Sharpe ratio of ca. 1.5. Note, as I preadjusted the return time series by its ex-ante market beta, this profile was achieved without deploying any significant market exposure. Second, all identified motifs $m_1, m_2, \dots, m_7$ exhibit risk-return profiles substantially better than the median period. The fact, that the motif discovery algorithm didn't produce any motifs attached to below-average periods is not surprising considering the factor's 'crashy' nature, i.e. its tendency to produce consistent returns in trendy states, but very harsh and fuzzy returns in reversion states.

Next, I calculate all motifs' risk-return profiles, rank them and assign signal values $s_{mx,y}$ accordingly. Exemplary, a factor return time series $T = t_1, t_2, \dots, t_n$ features two motifs m_1, m_2 that carry above-average risk-return profiles $p_{m1} > p_{m2} > \tilde{p}$ and two motifs m_3, m_4 with below-average profiles $p_{m4} < p_{m3} < \tilde{p}$, the algorithm assigns the signals $s_{m1} = 2, s_{m2} = 1, s_{m3} = -1, s_{m4} = -2$ throughout the time series. All periods t_{no} that do not feature a motif are assigned the signal value $s_{no} = 0$.

Note, that this algorithm might generate overlapping periods. In these instances, the economic regime identification model will use the most promising motif, i.e. the one that differs most severely – in absolute terms - from the median period[23]. Again using AQR's momentum factor between November 1995 and April 1996 as an example, the algorithm proceeds as follows:

Time	*Sharpe Ratio*	*Spread to Median*	*t+1*	*t+2*	*t+3*	*t+4*	*t+5*	*t+6*
Motif 1	4,0	2,5	0	0	0	0	0	0
Motif 2	3,9	2,4	4	4	4	4	4	4
Motif 3	3,3	1,8	0	0	0	0	0	3
Motif 4	4,1	2,6	0	6	6	6	6	6
Motif 5	2,8	1,3	0	0	0	0	0	0
Motif 6	5,0	3,5	0	0	7	7	7	7
Motif 7	1,8	0,3	0	0	0	0	0	0
	Composite Signal		4	6	7	7	7	7

Figure 10: Illustration of the model's processing of overlapping signals

Overall, I identify 28 motifs for the six AQR factors in the training dataset (1993 to 2011), 24 of which carry an above-average and four (2x Size, 1x Value FF, 1x Value AQR) a below-average risk-return profile[24]. I find the most motifs in Momentum and Value AQR (seven) and the least in Size (two). The following figure plots all signal time series identified

[23] I also considered z-scores and asymmetric transformations that attach more weight to deviations to the downside to account for the potential risk aversion of the investor, but finally settled on simple spreads. Given the substantial amount of parametrization required for the economic regime identification performed later, I do not regard the exact calibration of the dependent variable crucial for the overall success of the strategy.

in the training dataset. Note, given the adaptive nature of the motif discovery algorithm to the formation of new clusters, the list of stable motifs substantially expands throughout the exercise.

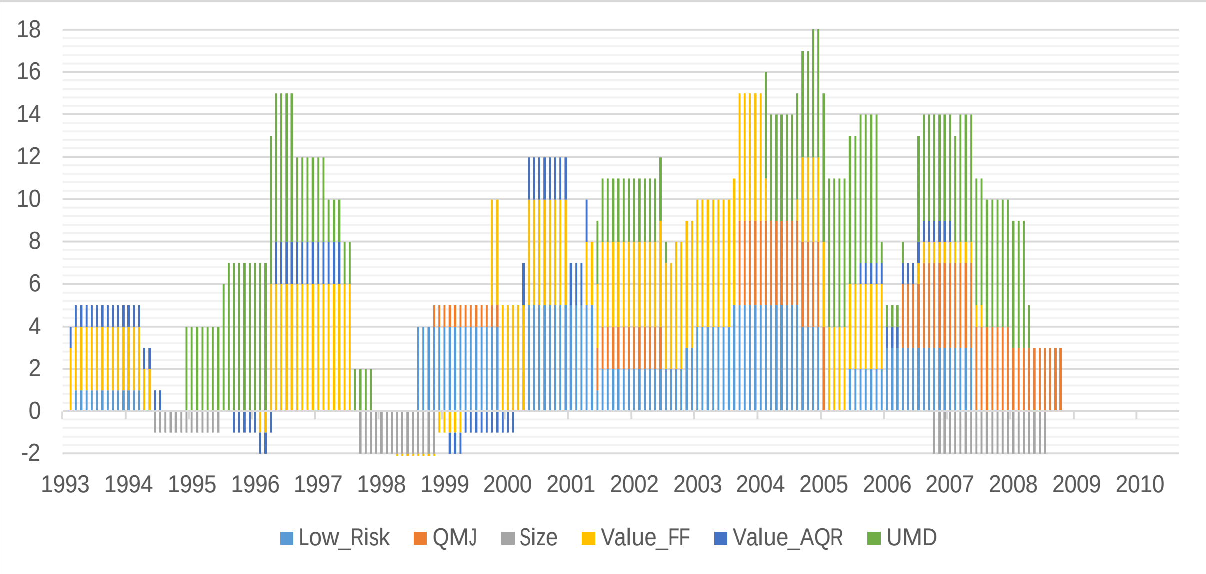

Figure 11: Identified motifs for AQR's six equity factors

Again, some interesting observations: The early 2000s all the way until the onset of the GFC were a great time to invest in factors; most of the signals overall as well as most of the strong signals are from this period. Second, the post-GFC period has so far produced no stable motif. However, given e.g. QMJ's and UMD's fairly consistent performance after 2010, I expect motifs to pop-up during the simulation of the trading strategy. Third, considering all factors and the corresponding motifs combined, the signal time series covers 89% of the entire sample. On an individual factor basis, however, the coverage ranges between 24% (Size) and 60% (Value FF). Lastly, the factors' motif overlap with another varies substantially across the different factor pairs. Exemplary, in 83% of periods in which Low-risk produces a signal, Value FF produces one as well. On the other hand, only in 7% of periods in which Value AQR produces a signal, Size produces one, too[25]. To reiterate, these overlaps do not pose a problem to either the economic regime identification algorithm or the trading strategy.

Summing-up, each factor's return time series is first clustered into several look-a-like subsequences and afterward converted to a signal time series $s_{mx,y}$ with the aforementioned $x = 28$ different motifs discovered in the training data set.

[25] The relative overlap figures of all pairs can be found in

Appendix C.

III.2.2 Economic Regime Identification

These signal time series $s_{mx,y}$ are then linked to a list of macro-economic indicators $E = e_1, e_2, \ldots, e_z$. As elaborated on in sections II.2 and III.1, no single macro-economic variable is a reliable predictor of either business cycle indicators (Stock & Watson, 2003) or factor returns (Hodges et al., 2017). Therefore, I adopt Bauer et al. (2004) and Arshanapalli et al. (2007) and linearly combine these individually unreliable forecasts via (ordered) logit regressions.

To avoid spurious findings resulting from potentially high serial correlation[26], however, I pre-process the data in three steps: First, I standardize every predictor variable e at time t by subtracting its arithmetic mean $\tilde{e}_i$ and dividing by its standard deviation σ_{e_i}. I use rolling windows covering the $N = 12$ months excluding t for the calculation of the mean and standard deviation (Ferson et al., 2003). Hence, the current observation of e is not included which allows for stronger innovations:

$$\text{(IX)} \quad e_t^{std} = \frac{e_t - \tilde{e}_i}{\sigma_{e_i}} = \frac{e_t - \frac{1}{N}\sum_{i=t-N}^{t-1} e_i}{\sqrt{\frac{1}{N-1}\sum_{i=t-N}^{t-1}(e_i - \bar{e})^2}}.$$

Second, I truncate all standardized predictor variables e_t^{std} at ± 5 to avoid extreme outliers (Dichtl et al., 2019). Lastly, I shrink the available set of predictor variables via principal component analysis (PCA) in the spirit of Neely et al. (2014). This additional pre-processing step is critical to avoid multi-collinearity problems within the regression analysis performed later. Moreover, the PCA eliminates unnecessary noise from the data set while consistently aggregating the heterogeneous information contained in the 65 predictor variables. While deriving the individual components is fairly straight-forward, settling on a specific sub-set is not. In general, there are three methods most popular in the data science community:

- Kaiser's rule, i.e. Eigenvalue > 1 (Kaiser, 1960)
- Scree or elbow plot rule (Cattell, 1966)
- Choose the amount of components c that they explain at least $th = [0{,}8; \ldots; 0{,}95; 0{,}99]$ of the cumulative variance

I also ran a PCA on three sub-sets (Valuation, business cycle and sentiment variables) of my 65 indicators and applied the three rules mentioned above as well as adapted Dichtl et al.

---

[26] Autocorrelation in macro-economic time series can be significant, as illustrated by the serial correlation of year-over-year relative change of the European Industrial Production Index: $\rho_{t,t-1} = 0{,}95$, $\rho_{t,t-2} = 0{,}89$, $\rho_{t,t-3} = 0{,}83$

(2019) by only using the first principal components of each sub-set. The following figure details the results for the training period 1993 to 2011:

Rule	*Joint Model*		*Valuation*		*Business Cycle*		*Sentiment*		*Combined*	
	#	% of Var	#	% of Var	#	% of Var	#	% of Var	#	# Difference
All	65	100%	8	100%	49	100%	8	100%	65	0
1st component	1	17%	1	45%	1	21%	1	32%	3	2
Kaiser's rule	16	76%	2	61%	12	75%	3	63%	17	1
Scree plot	22	84%	6	94%	21	88%	5	83%	32	10
th > 0,80	19	81%	4	83%	15	81%	5	83%	24	5
th > 0,85	24	86%	5	90%	19	86%	6	92%	30	6
th > 0,90	30	90%	6	94%	23	90%	6	92%	35	5
th > 0,95	40	95%	7	98%	31	95%	7	98%	45	5
th > 0,99	55	99%	8	100%	42	99%	8	100%	58	3
Used figure	23	85%	6	94%	20	87%	5	83%	31	8

Figure 12: Sensitivity tests of PCA results for the 65 macro-economic indicators for the period July 1993 to January 2011 Note, '#' denotes the number of components used and '% of Var' the share of cumulative variance explained using this number of components used. '# Difference' presents the difference between the sum of the number of components of the three individual sub-sets (Valuation, Business Cycle, Sentiment) and the Joint Model, which uses all 65 variables at once.

I round-up the averaged median and mean results for the Kaiser's, Scree plot and $th > 0{,}8$ rule to arrive at a 'best-guess' figure. Considering the eight 'excess' components required by the sum-of-parts process, I simply use the first 23 components of the joint dataset. They explain ca. 85% of the cumulative variance.

Given the granular nature of my motifs and the quantifiable difference in Sharpe ratio between two given motifs, I run ordered logit regression models to estimate the probability of an upcoming motif match. These models are set-up as follows:

Let $S_{mx,y}$ be an ordinal outcome with J possible realizations, then $P(S_{mx,y} \leq j)$ is the cumulative probability of $S_{mx,y}$ less than or equal to a specific realization $j = 1, \ldots, J-1$. The odds of being less than or equal a particular category can thus be defined as $\frac{P(S_{mx,y} \leq j)}{P(S_{mx,y} > j)}$ for $j = 1, \ldots, J-1$ since $P\left(S_{mx,y} > J\right) = 0$ and dividing by zero is undefined. Taking the log odds, also known as the logit, then results in

(X) $$log \frac{P(S_{mx,y} \leq j)}{P(S_{mx,y} > j)} = logit(P(Y \leq j)).$$

The ordinal logistic regression model thus is parameterized as

(XI) $$logit\left(P\left(S_{mx,y} \leq j\right)\right) = \beta_{j0} + \eta_1 x_1 + \cdots + \eta_{23} x_{23},$$

With $X = x_1, x_2, \dots x_{23}$ representing the 23 principal components, $\Theta = \eta_1, \eta_2, \dots, \eta_{23}$ their respective betas and β_{j0} the regression intercept. In short, ordered logit regressions extend the traditional logit regressions by allowing more-dimensional categories (or 'outcomes') in the analysis. On the other hand, ordered regressions require the categories to be in a natural order. Therefore, they represent a special case of the broader class of ordinal regression models.

The following figure indicates the results for the transformed momentum factor UMD. The output from the regression model for the first category, i.e. the motif associated with the smallest increase in Sharpe ratio vs. the median is shown for illustrative purposes.

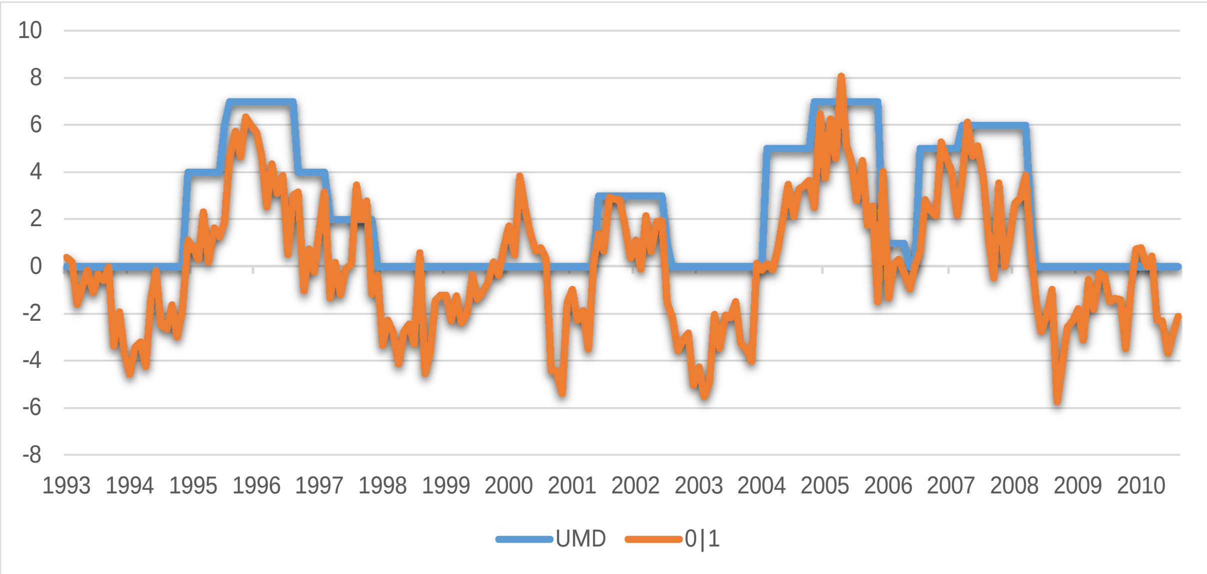

Figure 13: Visualization of the logit model results for the momentum factor's signal time series
Note: The regressions coefficients and their statistical significance for this example can be found in Appendix D.

Summing-up, ordered logit regression models are used to link the identified motifs to the dataset of macro-economic indicators. To avoid multi-collinearity issues, the dataset is further transformed using principal component analysis.

III.2.3 Trading Strategy

To recap, the main trading strategy of this paper is based on i) the successful identification of motifs (section III.2.1) as well as ii) the proper description of these periods using macro-economic indicators (section III.2.2). One way to identify the relative gain of incorporating motifs into a macro-economic 'matching' model via the aforementioned ordered logit algorithm is to compare the 'joint' performance to a simple model solely based on raw returns and macro-economic data.

Therefore, I first generate a static equal-weight portfolio that is rebalanced to its intended allocation every month. This portfolio serves as the 'static' portfolio. Second, I generate a simple trading strategy based on the list of indicators mentioned in section III.1. As classification models in the spirit of positive/ negative or above-average/ below-average typically outperform level estimation models (Leung et al., 2000), I transform the raw factor return time series as follows:

Given the return profile time series $P_j = p_{j,0}, p_{j,1}, \dots, p_{j,n}$ of factor j, I can calculate the standardized return profile time series $P_j^{std} = p_{j,0}^{std}, p_{j,1}^{std}, \dots, p_{j,n-12}^{std}$ by subtracting the series' mean $\widetilde{p_{j,\iota}}$ from every observation and dividing this term by the series' standard deviation $\sigma_{p_{j,i}}$. I use rolling windows covering the $N = 12$ months excluding t for the calculation of the mean and standard deviation. Hence, the current observation of e is not included which allows for stronger innovations:

$$\text{(XII)} \quad p_j^{std} = \frac{p_j - \widetilde{p_{j,\iota}}}{\sigma_{p_{j,i}}} = \frac{p_{j,i} - \frac{1}{N}\sum_{i=t-N}^{t-1} p_{j,k}}{\sqrt{\frac{1}{N-1}\sum_{i=t-N}^{t-1}(p_{j,i} - \widetilde{p_{j,\iota}})^2}}.$$

In turn, I assign each observation p_j^{std} a signal value according to the following matching algorithm.

$$\text{(XIII)} \quad s_j^{std} = \begin{cases} -2, if -3 < p_j^{std} \leq -2 \\ -1, if -2 < p_j^{std} \leq 0 \\ \dots \end{cases}.$$

Analog to the motif portfolio introduced next, this 'macro' portfolio is then linked to the macro-economic dataset using an ordered logit model. I calibrate the logit model using two-thirds of the available dataset - the remainder is used for the out-of-sample analysis. This 'macro' model is fully invested in the 'static' portfolio introduced above at all times. Given the signal from the model, however, the strategy adds or subtracts from this allocation. Thus, the 'macro' portfolio can be more or less significantly exposed to the underlying factor movements than the static portfolio. As all factors are dollar-neutral by definition, however, this should not materially affect the factor portfolio's cash balance[27].

Lastly, the main trading strategy of this paper matches the same reduced list of macro-economic indicators with another transformed factor return time series. This time, the transformation is done via the motif discovery algorithm introduced in section III.2.1. Analog to the 'macro' strategy, the 'motif' strategy is always invested in the static portfolio, too, but

[27] I do, however, report 'leverage'-adjusted figures in section IV. The conclusion remains unchanged.

can override this base portfolio allocation when the model generates a stable signal for one of the factors.

All trading strategies start on July 31st, 1993 and are reviewed monthly, i.e. all models are updated to include the most recent data and trading in and out of specific factors takes place. Note, as I calibrated the logit model on one month-lagged macro-economic indicators, the most recent observation can be used to predict the prospect of the upcoming period. Moreover, as this thesis focuses on high-conviction ideas, both the 'macro' and the 'motif' strategy only engage in the factor – long and short - if the model generated probability[28] of a positive (negative) period is larger than 0.75.

I calculate returns both including and excluding transaction costs. While many studies have dealt with the overall capacity of 'zero-investment' strategies (e.g. Ratcliffe et al., 2017), only a handful provide actual transaction cost estimates. Exemplary, Frazzini et al. (2018b) use transaction reports from AQR's massive database consisting of $1.7 trillion worth of trades and find average (median) price impact of 9.97 bp (6.18 bp) for executing the international Fama-French long-short factors between 1998 and 2016. The estimates for the more recent period 2006 to 2016 come in a bit lower at 9.54 bp (6.15 bp). Given that these estimates feature a fair share of trading in less liquid international markets like Argentina, Australia or Indonesia (Ng et al., 2015), these estimates are probably overstating the actual trading costs incurred by a factor strategy focused on European stocks.

Other studies calculate average annualized transaction costs and find somewhat similar results: Exemplary, Esakia et al. (2017) derive implementation costs for equal- and fundamental-weighted smart beta indices for US-based stocks between 1972 and 2014 between 11 bp to 42 bp, depending on the used universe's size as well as portfolio turnover.

In this paper, I assume 10 bp for one-way trading of incremental positions, 40 bp p.a. for maintaining the factor portfolio and 96 bp p.a. for holding the total return swaps (Dichtl et al., 2019). The difference between the first two figures arises because of the slow-moving nature of a lot of value, size or low-risk factor portfolios. Typically, these portfolios turn-over significantly less than once a year. On the other hand, momentum portfolios typically alter more often and thus require more trading to maintain the desired factor exposure. Because the unadjusted static strategy doesn't trade incrementally, it only incurs costs of 136 bp p.a. in 'maintenance' costs.

[28] Similar works often use 0.55 as the threshold (Bauer et al., 2004), but as this thesis focuses on a high-conviction tactical asset allocation framework, a more conservative approach seems warranted.

IV. Empirical Results

This section showcases the most important empirical results. As mentioned above, both the macro and the motif strategy tactically trade in and out of factors, in the process 'leveraging' the exposure to the underlying factor movements by a factor between -4 (0) and 5 (7) times for the macro (motif) portfolio. To report comparable figures, I thus scale the macro and the static portfolios according to the implicit leverage implemented in the motif portfolio.

Beginning with the risk-return profile calculated on a gross basis, the following figure illustrates the substantial risk-adjusted performance of the motif portfolio vs. the 'leverage'-adjusted macro and static portfolio.

	Macro Portfolio	*Motif Portfolio*	*Static Portfolio*
Mean Return p.a.	16,3%	41,2%	17,2%
CARR	13,0%	39,5%	16,8%
Annualized Volatility	24,3%	16,1%	7,9%
Ann. Downside Volatility	13,4%	8,5%	2,4%
Skewness	0,0	-1,2	0,7
Excess Kurtosis	1,5	8,3	0,4
Maximum Drawdown	-38,8%	-22,5%	-8,4%
Sharpe ratio	0,7	2,6	2,2
Sortino ratio	1,0	4,7	7,1
Omega ratio	0,3	1,7	2,0
MAR ratio	1,7	5,8	5,6

Figure 14: Risk-return profile of the three leverage-adjusted strategies, calculated on a gross of costs basis from January 2011 to November 2019

Overall, the motif strategy substantially outperforms the leverage-adjusted macro strategy in each dimension and also fares reasonably well versus the adjusted buy-and-hold strategy on a gross of costs basis: The Sharpe and MAR ratio improve, but both Sortino and Omega ratio deteriorate in relative terms. These findings are also robust to not adjusting the leverage.

To further rationalize this result, I run linear regressions on the three leverage-adjusted portfolio returns using the underlying size, value (2x), low-risk, quality (QMJ) and momentum (UMD) as dependent variables.

The following figure illustrates the findings:

Gross of costs		*Macro Portfolio*	*Motif Portfolio*	*Static Portfolio*
$Alpha_{monthly}$		-0,01	0,00	0,00
	t-stat	-1,43	-0,40	2,42
$Alpha_{annualized}$		-12,5%	-1,5%	4,0%
β_{Size}		1,20	-0,07	0,53
	t-stat	2,62	-0,38	6,70
$\beta_{Value\ AQR}$		0,75	0,91	0,95
	t-stat	0,57	1,73	4,16
$\beta_{Value\ FF}$		-0,94	0,07	0,21
	t-stat	-0,73	0,14	0,92
$\beta_{Low-risk}$		0,47	0,79	0,46
	t-stat	1,21	5,07	6,88
β_{QMJ}		1,11	1,26	0,42
	t-stat	1,57	4,42	3,41
β_{UMD}		0,96	1,30	0,61
	t-stat	1,79	6,03	6,60
Degrees of Freedom (df)		101	101	101
R-squared		26,2%	73,0%	79,3%

Figure 15: Regression results for leverage-adjusted portfolios vs. the six AQR factors Note: This table presents the results of the regressions of the total period returns (gross of cost) of the macro, motif and static portfolio. β_{Size}, $\beta_{Value\ AQR}$, $\beta_{Value\ FF}$, $\beta_{Low-risk}$, β_{QMJ}, β_{UMD} and represent the regression coefficients of the size, Value AQR, Value FF, Low-risk, quality and momentum factor. $Alpha_{monthly}$ represents the regression intercept,$Alpha_{annualized}$ is the annualized regression intercept. Below, I calculate t-statistics for each coefficient and intercept.

Note the less-than-perfect relationship between the static portfolio and its underlying equity factors. This is due to the leverage-adjustment to the motif portfolio. Put differently, by increasing the static portfolio's overall exposure to the underlying factors in good times and vice versa, the portfolio structure of the motif portfolio improves the risk-return profile of the static portfolio substantially, as evident by the significant alpha[29]. On its own, however, the motif portfolio doesn't add significant alpha. Yet, the motif strategy successfully picked its factor exposure: While the motif portfolio loads significantly on the three well-performing factors QMJ, UMD and Low-risk, its exposure to the underperforming factors size and value is insignificant.

[29] This effect is also visible in the static portfolio's loadings on its underlying factors. By definition, the betas of all factors to the static portfolio without leverage-adjustments equal the portfolio weight of $\frac{1}{6}$. Calculating the t-stats of the leverage-adjusted betas assuming $\beta_0 = \frac{1}{6}$, however, all factors except Value FF provide t-stats larger than two.

From an asset pricing model perspective, however, the motif portfolio does generate significant alphas. Regressing the six European Fama-French factors on the three portfolio returns, I derive the following estimates:

Gross of Costs		*Macro Portfolio*	*Motif Portfolio*	*Static Portfolio*
$Alpha_{monthly}$		0,00	0,01	0,01
	t-stat	0,30	3,84	4,53
$Alpha_{annualized}$		2,6%	19,2%	11,7%
β_{Mkt-Rf}		0,30	0,17	0,10
	t-stat	1,65	1,78	1,90
β_{SMB}		1,13	-0,11	0,61
	t-stat	2,62	-0,49	5,02
β_{HML}		-0,83	0,76	0,57
	t-stat	-1,22	2,09	2,96
β_{RMW}		-0,01	1,50	0,48
	t-stat	-0,01	3,45	2,09
β_{CMA}		0,56	0,15	0,44
	t-stat	0,71	0,35	1,98
β_{UMD}		0,70	1,09	0,32
	t-stat	2,56	7,47	4,12
Degrees of Freedom (df)		101	101	101
R-squared		18,7%	47,6%	39,4%

Figure 16: Regression results for the three leverage-adjusted portfolios vs. the six European Fama & French factors, all gross of costs. Note: This table presents the results of the regressions of the total period returns (gross of cost) of the macro, motif and static portfolio. β_{Mkt-Rf}, β_{SMB}, β_{HML}, β_{RMW}, β_{CMA}, β_{UMD} and represent the regression coefficients of the market, size, value, profitability, investment and momentum factor. $Alpha_{monthly}$ represents the regression intercept,$Alpha_{annualized}$ is the annualized regression intercept. Below, I calculate t-statistics for each coefficient and intercept.

Overall, all portfolio demonstrates insignificant betas to the market. Therefore, all strategies could be used as a TAA tool. The motif portfolio shows the best performance, however, with a significant alpha, strong betas to the 'good' factors UMD and RMW and limited exposure to the others. Yet, with R-squared figures between 0.19 and 0.48, the six-factor models do not convincingly explain the variance of the three portfolios.

Next, I analyze the performance net of costs. As elaborated on in section III.2.3, I assume 10 bp for one-way trading of incremental positions, 40 bp p.a. for maintaining an unlevered factor portfolio and 96 bp p.a. for holding the total return swaps (Dichtl et al., 2019). Overall, I derive total mean annualized costs of 3.98% (static portfolio), 4.18% (macro) and 6.33% (motif). Note, the differences between the total mean annualized costs and the return differential between the gross and net figures are due to annualization (for the mean) and the cumulative nature of the return time series (CARR).

The following figure visualizes the performance over the out-of-sample period:

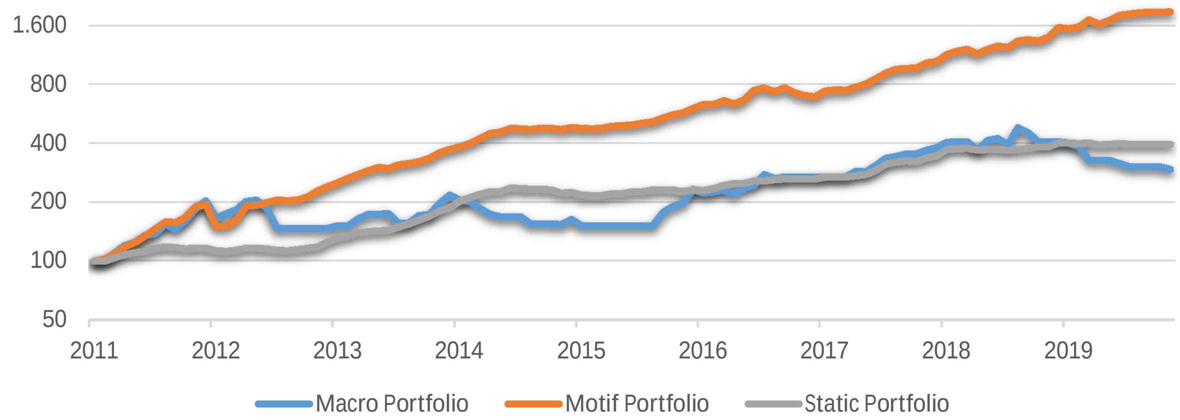

Figure 17: Visualization of leverage-adjusted performance of the three portfolios, net of costs.

Overall, all portfolios would have provided the investor with annualized returns around 10%. This number is particularly appealing as the profiles were generated without deploying any direct market exposure. Moreover, the macro portfolio's risk-return profile suffers from the material swings in its allocation. While it tracked the motif portfolio reasonably well until mid-2012, it re-allocated to weaker factors (e.g. Value FF) and underperformed afterward. Lastly, the static portfolio generated a very smooth return profile, with an annualized downside volatility of less than 3%. The following figure features all adjusted risk-return characteristics:

Net of Costs	*Macro Portfolio*	*Motif Portfolio*	*Static Portfolio*
Mean Return p.a.	11,7%	32,8%	12,7%
CARR	8,3%	31,1%	12,4%
Annualized Volatility	24,8%	16,0%	7,7%
Ann. Downside Volatility	14,6%	8,8%	2,7%
Skewness	-0,2	-1,2	0,6
Excess Kurtosis	1,5	8,6	0,4
Maximum Drawdown	-38,3%	-23,6%	-11,3%
Sharpe ratio	0,5	2,0	1,7
Sortino ratio	0,6	3,5	4,5
Omega ratio	0,2	1,3	1,1
MAR ratio	1,5	4,3	3,6

Figure 18: Risk-return profile of the three leverage-adjusted strategies, calculated on a net of costs basis from January 2011 to November 2019

In general, all strategies generate meaningful returns over the 2011 to 2019 period. They do so with different levels of risk, though: While the macro model experienced substantial drawdowns and downside volatility, the static model behaved more cautiously. The motif model, interestingly, exhibits both stretched negative skewness as well as high excess kurtosis. It still leads the other models according to the Omega ratio. Lastly, the motif portfolio's Sharpe ratio excels both the macro and the static portfolio's Sharpe ratios for one-way trading costs of up to 0.90%.

Delving deeper into the portfolio's return drivers, I again run linear regressions using Fama & French's (2018) six-factor model. As the factor loadings have not materially changed, I only report the intercept coefficient:

Net of Costs	*Macro Portfolio*	*Motif Portfolio*	*Static Portfolio*
$Alpha_{monthly}$	0,00	0,01	0,01
t-stat	-0,28	2,50	3,08
$Alpha_{annualized}$	-2,5%	12,1%	7,5%
Degrees of Freedom	101	101	101
R-squared	20,5%	48,0%	40,3%

Figure 19: Regression results for the three leverage-adjusted portfolios vs. the six European Fama & French factors, all net of costs. Note: This table presents the regressions results of the total period returns (net of cost) of the macro, motif and static portfolio. $Alpha_{monthly}$ represents the regression intercept,$Alpha_{annualized}$ is the annualized regression intercept. Below, I calculate t-statistics for each coefficient and intercept.

Overall, incorporating implementation costs subtracts ca. 7% p.a. (motif portfolio), 5% p.a. (macro) and 4% p.a. (static) of alpha. Yet, both the motif and the static portfolio still generate significant alpha.

To sum up, the motif portfolio outperforms both the macro and the static portfolio by a wide margin in raw and risk-adjusted terms. Moreover, in contrast to the macro portfolio, it also generates a six-factor alpha significant for one-way trading costs of up to 0.50%.

V. Discussion

Overall, the results of this work are encouraging: The motif model substantially outperformed both the simple macro model and a static equal-weight portfolio, before and after accounting for transaction costs. In contrast to the macro model, it also demonstrates a significant and exploitable alpha in the Fama & French (2018) six-factor model. There is a caveat, though. Data spanning essentially three boom (pre-2000, the early 2000s and post-2009) and two bust (Dotcom bubble, Global Financial Crisis) phases might not be adequate for a study dealing with the 'seasonality' of factor returns. In particular, as there were major differences in the factors driving the respective phases. Consider the post-Dotcom and post-GFC period: The former was clearly characterized by global housing and commodity booms – as evident in copper's rise from $1.500 to $8.5000 and the 30%-decline of the broad US Dollar index - accompanied by moderate but still positive interest rates and non-perfectly synchronized monetary fiscal policies. The second, however, was marked by extraordinary episodes of global policy convergence, through quantitative easing, zero-bound policy rates as well as synchronized fiscal stimuli. Lower and in most developed economies even negative real interest rates for an extended period, substantial compressions in yield spreads and multi-asset correlations at peak levels speak to that convergence as well (Bellone et al., 2019). Thus, longer-term studies covering hundred years of data also focusing on longer potential subsequences are needed to analyze these structural trends in more detail.

While there are some efforts explicitly looking at very long periods (e.g. Asness et al., 2013), also on a multi-asset basis, the task of analyzing the factors' time-varying premia and their relation with the macro-economic backdrop is still very much left to future research. In this context, the methodology of this paper can be further refined by e.g. incorporating additional transformations of the macro-economic indicators, choosing optimal lags according to AIC, BIC or a combination of several information criterions or applying support vector regressions instead of ordered logit regressions to tackle non-linear relationships between the return time series and the macro-economic indicators. It might also be interesting to test the subsequence time series clustering approach for other subsequence lengths like six or 24 months, to consistently apply dynamic time warping as the distance measure or incorporate implementation cost in the signal time series. Lastly, factor correlations could be taken into account by deploying multinomial models and risk management systems be developed that assist the manager to decide when to switch from a legacy to a more recent signal.

There is a lot of work to be done.

References

Aghabozorgi, S., Seyed Shirkhorshidi, A., & Ying Wah, T. (2015). Time-series clustering – A decade review. *Information Systems, 53*, 16-38.

Ahlborn, M., & Wortmann, M. (2018). The core–periphery pattern of European business cycles: A fuzzy clustering approach. *Journal of Macroeconomics*, *55*, 12-27.

Ahmed, P., Lockwood, L. J., & Nanda, S. (2002). Multistyle Rotation Strategies. *The Journal of Portfolio Management*, *28*(3), 17-29.

Ahn, S. C., Horenstein, A. R., & Wang, N. (2011). Determining the Rank of the Beta Matrix in Linear Asset Pricing Models. *SSRN Electronic Journal*.

Amenc, N., Malaise, P., Martellini, L., & Sfeir, D. (2003). Tactical Style Allocation—A New Form of Market Neutral Strategy. *The Journal of Alternative Investments, 6*(1), 8-22.

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31-56.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, *17*(2), 223-249.

Ammann, M., & Verhofen, M. (2006). The Effect of Market Regimes on Style Allocation. *Financial Markets and Portfolio Management*, *20*(3), 309-337.

Ang, A., Madhavan, A., & Sobczyk, A. (2017). Estimating time-varying factor exposures (Corrected October 2017). *Financial Analysts Journal*, *73*(4), 41-54.

Arnott, R. D., Beck, N., Kalesnik, V., & West, J. (2016). How can 'Smart beta' go horribly wrong? *Research Affiliates*.

Arshanapalli, B. G., Switzer, L. N., & Panju, K. (2007). Equity-style timing: A multi-style rotation model for the Russell large-cap and small-cap growth and value style indexes. *Journal of Asset Management, 8*(1), 9-23.

Asness, C. S., Frazzini, A., & Pedersen, L. H. (2018c). Quality minus junk. *Review of Accounting Studies*, *24*(1), 34-112.

Asness, C. S., Frazzini, A., Israel, R., & Moskowitz, T. J. (2015). Fact, fiction, and value investing. *The Journal of Portfolio Management*, *42*(1), 34-52.

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and Momentum Everywhere. *The Journal of Finance*, *68*(3), 929-985.

Asness, C., Chandra, S., Ilmanen, A., & Israel, R. (2017a). Contrarian Factor Timing is Deceptively Difficult. *The Journal of Portfolio Management*, *43*(5), 72-87.

Asness, C., Frazzini, A., Gormsen, N. J., & Pedersen, L. H. (2018b). Betting Against Correlation: Testing Theories of the Low-risk Effect. *Centre for Economic Policy Research*, *12686*.

Asness, C., Frazzini, A., Israel, R., Moskowitz, T. J., & Pedersen, L. H. (2018a). Size matters, if you control your junk. *Journal of Financial Economics*, *129*(3), 479-509.

Asness, C., Ilmanen, A., & Maloney, T. (2017b). Market Timing: Sin a Little Resolving the Valuation Timing Puzzle. *Journal Of Investment Management*, *15*(3), 23–40.

Bagnall, A., Lines, J., Bostrom, A., Large, J., & Keogh, E. (2017). The great time series classification bake off: A review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery*, *31*(3), 606-660.

Baker, M., & Wurgler, J. (2006). Investor Sentiment and the Cross-Section of Stock Returns. *The Journal of Finance*, *61*(4), 1645-1680.

Baker, M., Bradley, B., & Wurgler, J. (2011). Benchmarks as Limits to Arbitrage: Understanding the Low-Volatility Anomaly. *Financial Analysts Journal*, *67*(1), 40-54.

Ball, R., Gerakos, J., Linnainmaa, J. T., & Nikolaev, V. V. (2015). Deflating profitability. *Journal of Financial Economics*, *117*(2), 225-248.

Banz, R. W. (1981). The relationship between return and market value of common stocks. *Journal of Financial Economics, 9*(1), 3-18.

Barberis, N., & Huang, M. (2002). Mental accounting, loss aversion, and individual stock returns. *The Journal of Finance, 56*(4), 1247-1292.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*(3), 307-343.

Basu, S. (1977). Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis. *The Journal of Finance, 32*(3), 663-682.

Batista, G. E., Keogh, E. J., Tataw, O. M., & De Souza, V. M. (2013). CID: An efficient complexity-invariant distance for time series. *Data Mining and Knowledge Discovery, 28*(3), 634-669.

Bauer, R., Guenster, N., & Otten, R. (2004). Empirical evidence on corporate governance in Europe: The effect on stock returns, firm value and performance. *Journal of Asset Management, 5*(2), 91-104.

Beck, N., Hsu, J., Kalesnik, V., & Kostka, H. (2016). Will your factor deliver? An examination of factor robustness and implementation costs. *Financial Analysts Journal, 72*(5), 58-82.

Bellone, B., Declerck, P., Nordine, M., & Vy, T. (2019). Multi-Asset Style Factors Have Their Shining Moments. *SSRN Electronic Journal*.

Bender, J., & Wang, T. (2016). Can the whole be more than the sum of the parts? bottom-up versus top-down multifactor portfolio construction. *The Journal of Portfolio Management, 42*(5), 39-50.

Bender, J., Briand, R., Melas, D., Subramanian, R. A., & Subramanian, M. (2013). Deploying multi-factor index allocations in institutional portfolios. *MSCI Research Insight*, 339-363.

Bender, J., Sun, X., Thomas, R., & Zdorovtsov, V. (2018). The Promises and Pitfalls of Factor Timing. *The Journal of Portfolio Management*, *44*(4), 79-92.

Berndt, O. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. *In Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining (AAAIWS'94). AAAI Press*, 359-370.

Bevelin, P. (2005). *Seeking Wisdom: From Darwin to Munger* (3rd ed.). PCA Publications L.L.c.

Beveratos, A., Bouchaud, J., Ciliberti, S., Laloux, L., Lempérière, Y., Potters, M., & Simon, G. (2014). Deconstructing the Low-Vol Anomaly. *The Journal of Portfolio Management Fall*, *44*(1), 91-103.

Bird, R., & Casavecchia, L. (2007). Value enhancement using momentum indicators: The European experience. *International Journal of Managerial Finance*, *3*(3), 229-262.

Black, F. (1972). Capital market equilibrium with restricted borrowing. *The Journal of Business*, *45*(3), 444.

Blitz, D. (2016). The value of low volatility. *Journal of Portfolio Management*, *42*(3), 94-100.

Blitz, D. (2017). Are Exchange-Traded Funds Harvesting Factor Premiums? *Journal of Investment Consulting*, *18*(1).

Bouchaud, J., Ciliberti, S., Landier, A., Simon, G., & Thesmar, D. (2016). The Excess Returns of 'Quality' Stocks: A Behavioral Anomaly. *HEC Paris Research Paper*, FIN-2016-1134.

Boudoukh, J., Michaely, R., Richardson, M., & Roberts, M. R. (2007). On the Importance of Measuring Payout Yield: Implications for Empirical Asset Pricing. *The Journal of Finance, 62*(2), 877-915.

Brandt, M. W., & Santa-Clara, P. (2006). Dynamic Portfolio Selection by Augmenting the Asset Space. *The Journal of Finance*, *61*(5), 2187-2217.

Brandt, M. W., Santa-Clara, P., & Valkanov, R. (2009). Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns. *Review of Financial Studies*, *22*(9), 3411-3447.

Brunnermeier, M. K., & Pedersen, L. H. (2009). Market Liquidity and Funding Liquidity. *The Review of Financial Studies*, *22*(6), 2201–2238.

Campbell, J. Y., & Shiller, R. J. (2001). Valuation Ratios and the Long-Run Stock Market Outlook. *The Journal of Portfolio Management*, *24(2), 11-26.*

Campbell, J. Y., & Vuolteenaho, T. (2004). Bad Beta, Good Beta. *The American Economic Review*, *94*(5), 1249-1275.

Campbell, J. Y., Hilscher, J., & Szilagyi, J. (2008). In Search of Distress Risk. *The Journal of Finance*, *63*(6), 2899-2939.

Campbell, J. Y., Polk, C., & Vuolteenaho, T. (2009). Growth or Glamour? Fundamentals and Systematic Risk in Stock Returns. *The Review of Financial Studies, 23*(1), 305–344.

Cattell, R. B. (1966). The Scree Test For The Number Of Factors. *Multivariate Behavioral Research*, *1*(2), 245-276.

Cecchetti, S. G., Kashyap, A. K., & Wilcox, D. W. (1997). Interactions Between the Seasonal and Business Cycles in Production and Inventories. *The American Economic Review*, *87*(5), 884-892.

Cenedese, G., Payne, R., Sarno, L., & Valente, G. (2016). What Do Stock Markets Tell Us about Exchange Rates? *Review of Finance*, *20*(3), 1045-1080.

Cenesizoglu, T., Liu, Q., Reeves, J. J., & Wu, H. (2016). Monthly Beta Forecasting with Low-, Medium- and High-Frequency Stock Returns. *Journal of Forecasting*, *35*(6), 528-541.

Chaieb, I., Langlois, H., & Scaillet, O. (2018). Time-Varying Risk Premia in Large International Equity Markets, (Revised Jun 2018). Swiss Finance Institute Research Paper Series 18-04. Swiss Finance Institute.

Chan, K., & Chen, N. (1991). Structural and Return Characteristics of Small and Large Firms. *The Journal of Finance*, *46*(4), 1467-1484.

Chan, K., Chan, L. K., Jegadeesh, N., & Lakonishok, J. (2006). Earnings Quality and Stock Returns. *The Journal of Business*, *79*(3), 1041-1082.

Chan, K., Chen, N., & Hsieh, D. A. (1985). An exploratory investigation of the firm size effect. *Journal of Financial Economics*, *14*(3), 451-471.

Chauvet, M., & Potter, S. (2000). Coincident and leading indicators of the stock market. *Journal of Empirical Finance*, *7*(1), 87-111.

Chen, N., & Zhang, F. (1998). Risk and Return of Value Stocks. *The Journal of Business*, *71*(4), 501-535.

Ciliberti, S., Sérié, E., Simon, G., Lempérière, Y., & Bouchaud, J. (2019). The Size Premium in Equity Markets: Where Is the Risk? *The Journal of Portfolio Management*, *45*(5), 58-68.

Clare, A., Sapuric, S., & Todorovic, N. (2010). Quantitative or momentum-based multi-style rotation? UK experience. *Journal of Asset Management*, *10*(6), 370-381.

Cochrane, J. H. (1991). Production-Based Asset Pricing and the Link Between Stock Returns and Economic Fluctuations. *The Journal of Finance*, *46*(1), 209-237.

Connor, G. (1995). The three types of factor models: A comparison of their explanatory power. *Financial Analysts Journal*, *51*(3), 42-46.

Conrad, J., Cooper, M., & Kaul, G. (2003). Value versus Glamour. *The Journal of Finance, 58*(5), 1969-1995.

Cooper, M. J., Gulen, H., & Schill, M. J. (2008). Asset Growth and the Cross-Section of Stock Returns. *The Journal of Finance*, *63*(4), 1609-1651.

Cooper, M. J., Gulen, H., & Vassalou, M. (2001). Investing in size and book-to-market portfolios using information about the macroeconomy: Some new trading rules. *SSRN Electronic Journal*.

Copeland, M. M., & Copeland, T. E. (1999). Market Timing: Style and Size Rotation Using the VIX. *Financial Analysts Journal*, *55*(2), 73-81.

Daniel, K. D., & Moskowitz, T. J. (2016). Momentum Crashes. *Journal of Financial Economics*, *122*, 221–247.

Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor Psychology and Security Market Under- and Overreactions. *The Journal of Finance*, *53*(6), 1839-1885.

Dasgupta, A., Prat, A., & Verardo, M. (2011). The price impact of institutional herding. *Review of Financial Studies*, *24*(3), 892-925.

Dechow, P., Ge, W., & Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics*, *50*(2-3), 344-401.

Dichtl, H., Drobetz, W., Lohre, H., Rother, C., & Vosskamp, P. (2019). Optimal Timing and Tilting of Equity Factors. *Financial Analysts Journal*, *75*(4), 84-102.

Driver, H. E., & Kroeber, A. L. (1932). Quantitative Expression of Cultural Relationships. *University of California Publications in American Archaeology and Ethnology*, 211-256.

Elsby, M. W., Shin, D., & Solon, G. (2016). Wage Adjustment in the Great Recession and Other Downturns: Evidence from the United States and Great Britain. *Journal of Labor Economics*, *34*(S1), 249-291.

Esakia, M., Goltz, F., Sivasubramanian, S., & Ulahel, J. (2017). Smart Beta Replication Costs. *EDHEC-Risk Institute Publication - February 2017*.

European Central Bank (ECB). (2011). *Monthly Bulletin May 2011*.

Faloutsos, C., Ranganathan, M., & Manolopoulos, Y. (1994). Fast Subsequence Matching in Time-Series Databases. *ACM SIGMOD Record*, *23*(2), 419-429.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, *25*(1), 23-49.

Fama, E. F., & French, K. R. (1992). The Cross-Section of Expected Stock Returns. *The Journal of Finance*, *47*(2), 427-465.

Fama, E. F., & French, K. R. (1995). Size and Book-to-Market Factors in Earnings and Returns. *The Journal of Finance*, *50*(1), 131-155.

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

Fama, E. F., & French, K. R. (2016). Dissecting Anomalies with a Five-Factor Model. *The Review of Financial Studies*, *29*(1), 69–103.

Fama, E. F., & French, K. R. (2017). International tests of a five-factor asset pricing model. *Journal of Financial Economics*, *123*(3), 441-463.

Fama, E. F., & French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, *128*(2), 234-252.

Ferson, W. E., Sarkissian, S., & Simin, T. T. (2003). Is Stock Return Predictability Spurious? *Journal of Investment Management*, *1*(3), 1-10.

Forina, M., Casale, M., & Oliveri, P. (2009). Application of Chemometrics to food chemistry. *Comprehensive Chemometrics*, 75-128.

Frazzini, A., & Pedersen, L. H. (2014). Betting against beta. *Journal of Financial Economics, 111*(1), 1-25.

Frazzini, A., Israel, R., & Moskowitz, T. J. (2018b). Trading Costs. *SSRN Electronic Journal*.

Frazzini, A., Kabiller, D., & Pedersen, L. H. (2018a). Buffett's Alpha. *Financial Analysts Journal*, *74*(4), 35-55.

Graham, B., & Dodd, D. L. (1934). *Security analysis: The classic 1951 edition*. Whittlesey House, McGraw-Hill Book Co..

Grobys, K., & Kolari, J. W. (2019). Another look on choosing factors: The international evidence. *SSRN Electronic Journal*.

Gupta, T., & Kelly, B. (2019). Factor Momentum Everywhere. *The Journal of Portfolio Management*, *45*(3), 13-36.

Haddad, V., Kozak, S., & Santosh, S. (2020). Factor timing. *The Review of Financial Studies*, *33*(5), 1980–2018.

Hameed, A., Kang, W., & Viswanathan, S. (2010). Stock Market Declines and Liquidity. *The Journal of Finance*, *65*(1), 257-293.

Hammerschmid, R., & Lohre, H. (2017). Regime Shifts and Stock Return Predictability. *International Review of Economics & Finance*, *56*, 138-160.

Harvey, C. R. (2017). Presidential address: The scientific outlook in financial economics. *The Journal of Finance*, *72*(4), 1399-1440.

Harvey, C. R., Liu, Y., & Zhu, H. (2016). . . . and the Cross-Section of Expected Returns. *The Review of Financial Studies*, *29*(1), 5-68.

Haugen, R. A., & Heins, A. J. (1975). Risk and the Rate of Return on Financial Assets: Some Old Wine in New Bottles. *The Journal of Financial and Quantitative Analysis, 10*(5), 775.

Hirshleifer, D., Hou, K., Teoh, S. H., & Zhang, Y. (2004). Do Investors Overvalue Firms with Bloated Balance Sheets? *Journal of Accounting and Economics, 38*(1), 297-331.

Hodges, P., Hogan, K., Peterson, J. R., & Ang, A. (2017). Factor Timing with Cross-Sectional and Time-Series Predictors. *The Journal of Portfolio Management, 44*(1), 30-43.

Hodrick, R. J. (1992). Dividend Yields and Expected Stock Returns: Alternative Procedures for Inference and Measurement. *The Review of Financial Studies, 5*(3), 357–386.

Hong, H., Lim, T., & Stein, J. C. (2000). Bad News Travels Slowly: Size, Analyst Coverage, and the Profitability of Momentum Strategies. *The Journal of Finance, 55*(1), 265-295.

Hou, K., & Moskowitz, T. J. (2005). Market Frictions, Price Delay, and the Cross-Section of Expected Returns. *The Review of Financial Studies, 18*(3), 981-1020.

Hou, K., Xue, C., & Zhang, L. (2015). Digesting Anomalies: An Investment Approach. *The Review of Financial Studies, 28*(3), 650-705.

Hsu, J. C., Kalesnik, V., & Kose, E. (2019). What Is Quality? *Financial Analysts Journal, 75*(2), 44-61.

Hsu, J. C., Kudoh, H., & Yamada, T. (2013). When Sell-Side Analysts Meet High-Volatility Stocks: An Alternative Explanation for the Low-Volatility Puzzle. *Journal Of Investment Management, 11*(2), 28-46.

Hsu, J., Kalesnik, V., & Viswanathan, V. (2015). A framework for assessing factors and implementing smart beta strategies. *The Journal of Index Investing, 6*(1), 89-97.

Ilmanen, A. S., Israel, R., Moskowitz, T. J., Thapar, A. K., & Wang, F. (2019). Factor premia and factor timing: A century of evidence. *SSRN Electronic Journal*.

Jacobs, B. I., & Levy, K. N. (1996). High-Definition Style Rotation. *The Journal of Investing*, *5*(3), 14-23.

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

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*(1), 141-151.

Kao, D., & Shumaker, R. D. (1999). Equity Style Timing. *Financial Analysts Journal*, *55*(1), 37-48.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics*, *17*(2), 357-390.

Keimling, N. (2016). Predicting Stock Market Returns Using the Shiller CAPE — An Improvement Towards Traditional Value Indicators? *StarCapital AG*.

Keogh, E., & Lin, J. (2005). Clustering of time-series subsequences is meaningless: implications for previous and future research. *Knowledge and Information Systems*, *8*(2), 154-177.

Keogh, E., & Ratanamahatana, C. A. (2005). Exact indexing of dynamic time warping. *Knowledge and Information Systems*, *7*(3), 358-386.

Koch, A., Ruenzi, S., & Starks, L. (2016). Commonality in Liquidity: A Demand-Side Explanation. *The Review of Financial Studies, 29*(8), 1943-1974.

Kondor, P., & Vayanos, D. (2019). Liquidity Risk and the Dynamics of Arbitrage Capital. *The Journal of Financial*, *74*(3), 1139-1173.

Koopman, S. J., Lit, R., & Lucas, A. (2017). Model-Based Business Cycle and Financial Cycle Decomposition for Europe and the U.S. *Systemic Risk Tomography*, 151-168.

Kose, M. A., Otrok, C., & Whiteman, C. H. (2003). International Business Cycles: World, Region, and Country-Specific Factors. *American Economic Review*, *93*(4), 1216-1239.

La Porta, R. (1996). Expectations and the cross-section of stock returns. *The Journal of Finance*, *51*(5), 1715-1742.

La Porta, R., Lakonishok, J., Shleifer, A., & Vishny, R. (1997). Good News for Value Stocks: Further Evidence on Market Efficiency. *The Journal of Finance*, *52*(2), 859-874.

Lakonishok, J., Shleifer, A., & Vishny, R. W. (1993). Contrarian Investment, Extrapolation, and Risk. *The Journal of Finance*, *49*(5), 1541-1578.

Leippold, M., & Rüegg, R. (2019). Fama–French Factor Timing: The Long-Only Integrated Approach. *SSRN Electronic Journal*.

Leivo, T. H., & Pätäri, E. J. (2011). Enhancement of value portfolio performance using momentum and the long-short strategy: The Finnish evidence. *Journal of Asset Management*, *11*(6), 401-416.

Leung, M. T., Daouk, H., & Chen, A. (2000). Forecasting stock indices: a comparison of classification and level estimation models. *International Journal of Forecasting*, *16*(2), 173-190.

Levis, M., & Liodakis, M. (1999). The Profitability of Style Rotation Strategies in the United Kingdom. *The Journal of Portfolio Management*, *26*(1), 73-86.

Levis, M., & Tessaromatis, N. (2004). Style Rotation Strategies. *The Journal of Portfolio Management*, *30*(4), 160-169.

Levy, R. A. (1967). Relative Strength as a Criterion for Investment Selection. *The Journal of Finance*, *22*(4), 595-610.

Li, H., Novy-Marx, R., & Velikov, M. (2019). Liquidity Risk and Asset Pricing. *Critical Finance Review*, *8*(1-2), 223-255.

Lochstoer, L. A., & Tetlock, P. C. (2020). What Drives Anomaly Returns? *The Journal of Finance, Forthcoming*.

López-Salido, D., Stein, J. C., & Zakrajšek, E. (2017). Credit-Market Sentiment and the Business Cycle. *The Quarterly Journal of Economics*, *132*(3), 1373-1426.

Loughran, T., & Ritter, J. R. (1995). The New Issues Puzzle. *The Journal of Finance*, *50*(1), 23-51.

Lucas, A., Van Dijk, R., & Kloek, T. (2002). Stock selection, style rotation, and risk. *Journal of Empirical Finance*, *9*(1), 1-34.

Martin, P. G., & McCann, B. B. (1989). *The investor's guide to fidelity funds*. John Wiley & Sons.

McLean, R. D., & Pontiff, J. (2016). Does Academic Research Destroy Stock Return Predictability? *The Journal of Finance*, *71*(1), 5-32.

Merton, R. C. (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*, *41*(5), 867-887.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *The Journal of Finance*, *29*(2), 449-470.

Miller, K. L., Li, H., Zhou, T. G., & Giamouridis, D. (2015). A risk-oriented model for factor timing decisions. *The Journal of Portfolio Management*, *41*(3), 46-58.

Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics*, *58*(1-2), 215-260.

Nalbantov, G., Bauer, R., & Sprinkhuizen-Kuyper, I. (2006). Equity Style Timing Using Support Vector Regressions. *Applied Financial Economics*, *16*(15), 1095-1111.

Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the Equity Risk Premium: The Role of Technical Indicators. *Management Science*, *60*(7), 1772-1791.

Ng, L., Wu, F., Yu, J., & Zhang, B. (2015). Foreign investor heterogeneity and stock liquidity around the world*. *Review of Finance*, *20*(5), 1867-1910.

Novy-Marx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, *108*(1), 1-28.

Novy-Marx, R., & Velikov, M. (2015). A Taxonomy of Anomalies and Their Trading Costs. *The Review of Financial Studies*, *29*(1), 104-147.

Parker, A. S., Hayes, B. T., Ortega, A., & Naha, C. (2016). The 2016 Factor Timing Quantroversy. *Morgan Stanley US Quant Research*.

Pástor, L., & Stambaugh, R. F. (2003). Liquidity Risk and Expected Stock Returns. *Journal of Political Economy*, *111*(3), 642-685.

Pástor, L., & Stambaugh, R. F. (2019). Liquidity Risk After 20 Years. *Chicago Booth Research Paper No. 19-13, Fama-Miller Working Paper*.

Petkova, R. (2006). Do the Fama-French factors proxy for innovations in predictive variables? *The Journal of Finance*, *61*(2), 581-612.

Pontiff, Y., & Singla, R. (2019). Liquidity Risk? *Critical Finance Review*, *8*(1-2), 257-276.

Rai, p., & Singh, S. (2010). A Survey of Clustering Techniques. *International Journal of Computer Applications*, *7*(12).

Rapach, D. E., Strauss, J. K., & Zhou, G. (2013). International Stock Return Predictability: What Is the Role of the United States? *The Journal of Finance*, *68*(4), 1633-1662.

Ratcliffe, R., Miranda, P., & Ang, A. (2017). Capacity of Smart Beta Strategies: A Transaction Cost Perspective. *Columbia Business School Research Paper*, 16-76.

Reinganum, M. R. (1981). Misspecification of capital asset pricing: Empirical anomalies based on earnings' yields and market values. *Journal of Financial Economics*, *9*(1), 19-46.

Rodpongpun, S., Niennattrakul, V., & Ratanamahatana, C. A. (2012). Selective Subsequence Time Series clustering. *Knowledge-Based Systems*, *35*, 361-368.

Roll, R. (1986). The Hubris Hypothesis of Corporate Takeovers. *The Journal of Business*, *59*(2), 197-216.

Rothman, M. S. (2007). U.S. Equity Quantitative Strategies. *Lehmann Brothers Equity Research*.

Shadwick, W. F., & Keating, C. (2002). A Universal Performance Measure. *Journal of Performance Measurement*, *6*(3), 59-84.

Stock, J. H., & Watson, M. W. (2003). Forecasting Output and Inflation: The Role of Asset Prices. *Journal of Economic Literature, 47*, 788-829.

Subbiah, M., & Fabozzi, F. J. (2016). Equity style allocation: A nonparametric approach. *Journal of Asset Management*, *17*(3), 141-164.

Tan, C. W., Webb, G. I., & Petitjean, F. (2017). Indexing and classifying gigabytes of time series under time warping. *Proceedings of the 2017 SIAM International Conference on Data Mining, 29.*

Uhl, M. W., Pedersen, M., & Malitius, O. (2015). What's in the News? Using News Sentiment Momentum for Tactical Asset Allocation. *The Journal of Portfolio Management, 41*(2), 100-112.

Vassalou, M., & Xing, Y. (2004). Default Risk in Equity Returns. *The Journal of Finance, 59*(2), 831-868.

Wang, X., Mueen, A., Ding, H., Trajcevski, G., Scheuermann, P., & Keogh, E. (2012). Experimental comparison of representation methods and distance measures for time series data. *Data Mining and Knowledge Discovery, 26*(2), 275-309.

Winkelmann, K., Suryanarayanan, R., Hentschel, L., & Varga, K. (2013). Macro-Sensitive Portfolio Strategies. *MSCI Market Insight.*

Yeh, C. M., Zhu, Y., Ulanova, L., Begum, N., Ding, Y., Dau, H. A., Silva, D. F., Mueen, A., & Keogh, E. (2016). Matrix Profile I: All Pairs Similarity Joins for Time Series: A Unifying View That Includes Motifs, Discords and Shapelets," 2016 IEEE 16th International Conference on Data Mining (ICDM), Barcelona, 2016, pp. 1317-1322.
XX

Zolhavarieh, S. J., Aghabozorgi, S., & Teh, Y. W. (2014). A Review of Subsequence Time Series Clustering. *The Scientific World Journal*.

Appendix A

Obtained factor return time series

As mentioned in section III.1, I use the available European return factor dataset available from AQR's website[30]. It features factor estimates for size (Size), quality (QMJ), low-risk (Low-Risk), momentum (UMD) and two for value (Value FF, Value AQR). In general, the factors are based on 17 European countries and are calculated in excess of the risk-free rate in USD-equivalents. The 17 European countries are: Austria, Belgium, Switzerland, Germany, Denmark, Spain, Finland, France, United Kingdom, Greece, Ireland, Israel, Italy, the Netherlands, Norway, Portugal and Sweden.

Appendix B

Obtained macro-economic indicators

I pulled the 65 indicators from a variety of sources, including Barclays, Bloomberg, Datastream, the FRED Economic Database as well as Global Financial Data (GFD). Moreover, I used data from the website of Prof Shiller[31] and Prof Goyal[32]. All variables are lagged by one month, and subsequently transformed via taking the (log) change over a 12 month basis. Past stock market returns, however, are calculated based on the past 6 months return. The indicators are the following:

Type	Europe	USA	Japan	Global	Comment
Industrial Production	Yes	Yes	Yes	No	From Datastream
Consumer Confidence	Yes	Yes	Yes	No	From Bloomberg
CPI Price Inflation	Yes	Yes	Yes	No	From Bloomberg
Real Money M1 Supply	Yes	Yes	Yes	No	Used 12 month change in CPI price inflation to deflate the

[30] Link: https://www.aqr.com/library/data-sets

31 Link. http://www.econ.yale.edu/~shiller/data.htm

32 Link: http://www.hec.unil.ch/agoyal/

					available nominal supply of M1 money
Currencies	EUR-USD, REER EUR	USD-GBP, REER USD	USD-JPY	No	From Bloomberg
Government Bonds	3m, 2y, 10y	3m, 2y, 10y	3m, 2y, 10y	No	3m Europe not available from 1992 onward, estimated based on a linear regression model using Italy, Spain, Germany, France and the UK as input
Interbank Rates	3m	3m	3m	No	Euribor only viable after 1999, used adjusted Fibor time series before
Corporate Bonds	IG	IG, HY, BBB – AAA	No	No	From Barclays
Stock Indices	Stoxx 600	S&P 500	Topix	No	From Bloomberg
Price of risk	No	Yes	No	No	Calculated as Vix Index divided by past 30d volatility of the S&P 500, data from Bloomberg

Valuations	Earnings Yield, Dividend Yield	Earnings Yield, Dividend Yield, Price/ Book, Shiller CAPE	Earnings Yield, Dividend Yield	No	Earnings and dividend yield data from GFD, CAPE from Prof Shiller's website, price/ book from Prof. Goyal's website
Unemployment Rate	Yes	Yes	Yes	No	From Bloomberg
Business Surveys	Ifo, Insee	ISM	Bank of Japan	No	Global PMIs not publicly available before 2015, all data from Bloomberg
Commodities	No	No	No	Gold, Copper, WTI Brent Oil, the IMF Commodity Price Index	IMF Commodity Price Index not available before 1995, thus proxied via Bloomberg's commodity Index BCOM before, all data from Bloomberg
Credit	No	Bank Credit, All Commercial Banks	No	No	From FRED database

HY issuance share	No	Yes	No	Yes	From Bloomberg’s SRCH function

Appendix C

Overlap of all factor pairs' signal time series

The following figure details the overlap of all factor pairs' time series. Note, the figures are not simply mirrored at the diagonal. Exemplary, in 83% of periods in which Low-risk produces a signal, Value FF produces one as well. On the other hand, only in 71% of periods in which Value FF produces a signal, Low-risk produces one, too.

	Low-Risk	*QMJ*	*Size*	*Value FF*	*Value AQR*	*UMD*
Low-Risk		76%	24%	71%	64%	47%
QMJ	51%		46%	36%	24%	48%
Size	11%	31%		14%	7%	29%
Value FF	83%	62%	36%		75%	64%
Value AQR	49%	27%	12%	49%		37%
UMD	43%	65%	58%	50%	45%	

Figure 20: Overlap of all factor pairs' signal time series

Appendix D

Regression coefficients of ordered logit regression

The following model was estimated using the following equation for the period of July 1993 to January 2011:

(XIV) $logit\big(P(UMD \leq j)\big) = \beta_{j0} - \eta_1 PC_1 - \cdots - \eta_{23} PC_{23}$, with $0 \leq j \leq 7$

Parameter	*Value*	*Std. Error*	*t value*	*p-value*
0/1	0,15	0,22	0,68	0,50
1/2	0,42	0,22	1,87	0,06
2/3	0,89	0,23	3,79	0,00
3/4	1,50	0,25	5,94	0,00
4/5	2,15	0,28	7,74	0,00
5/6	3,14	0,33	9,55	0,00
6/7	4,09	0,39	10,62	0,00
PC1	-0,06	0,06	-1,01	0,31
PC2	0,00	0,06	0,02	0,98
PC3	-0,03	0,08	-0,36	0,72
PC4	-0,35	0,08	-4,33	0,00
PC5	-0,41	0,09	-4,41	0,00
PC6	-0,43	0,10	-4,17	0,00
PC7	-0,71	0,12	-5,83	0,00
PC8	0,07	0,11	0,61	0,54
PC9	-0,77	0,15	-5,21	0,00
PC10	0,52	0,13	4,12	0,00
PC11	0,81	0,15	5,50	0,00
PC12	-0,54	0,16	-3,43	0,00
PC13	0,24	0,14	1,69	0,09
PC14	-0,58	0,15	-3,91	0,00
PC15	-0,04	0,17	-0,25	0,80
PC16	-0,48	0,16	-2,93	0,00
PC17	1,03	0,19	5,51	0,00
PC18	0,29	0,20	1,45	0,15
PC19	0,05	0,17	0,28	0,78
PC20	0,00	0,19	0,02	0,98
PC21	0,03	0,20	0,17	0,86
PC22	-0,66	0,22	-3,04	0,00
PC23	0,04	0,22	0,20	0,84

Figure 21: Regression coefficients of ordered logit regression