A Work Project, presented as part of the requirements for the Award of a Master Degree in Finance from the NOVA – School of Business and Economics.

Multi Factor Investing: A Study on Market Efficiency Evolution

Tiago Correia (34084)

A Project carried out on the Master in Finance Program, under the supervision of

Gonçalo Sommer Ribeiro

January 3, 2020

Abstract

This paper was built with the intention of verifying the evolution, over the last decades, of the statistical significance of traditional factors used for equity factor investing, such as SMB (small minus big), HML (high minus low), MOM (momentum) and MKRTF (market risk premium), and others. The results obtained after the construction of a daily multi-factor portfolio and several single factor portfolios concluded on the rejection of the hypothesis that these factors are statistically significant, both globally and individually, even though financial results are, in some periods, strong enough to consider these strategies reliable for trading purposes.

Keywords: Efficient Market Hypothesis, Factor Investing, Long-short strategies, Transaction costs

This work used infrastructure and resources funded by Fundação para a Ciência e a Tecnologia (UID/ECO/00124/2013, UID/ECO/00124/2019 and Social Sciences DataLab, Project 22209), POR Lisboa (LISBOA-01-0145-FEDER-007722 and Social Sciences DataLab, Project 22209) and POR Norte (Social Sciences DataLab, Project 22209).

1. Introduction

Equity investors have been for decades trying to find ways to outperform the market by incessantly searching for reliable strategies, as factor-based investing, moving averages, fundamental valuation, etc... This hunt for systematic higher yields led to the study of the relationship between risk and return, influencing the creation of models such as the Capital Asset Pricing Model by Treynor (1961, 1962) and Sharpe (1964), which forecasts a positive linear relation between risk and return.

Even though factor-based investment, the selection of securities based on certain characteristics to achieve improved returns, has been on the rise over the years, it has been pronounced by Hunstad and Dekhayser (2015) as no longer a reliable investment strategy. This was due to the fact that the alpha it used to generate for many factors has disappeared. A possible cause for this disappearance might be, for example, that investors have now more information available for companies that used to be ignored for trading purposes (companies with small market capitalization). So far, around 300 factors have been identified and published so far (Harvey and Liu, 2015) as factors capable of generating extraordinary returns, but when tested for their statistical significance the majority resulted in statistical insignificance. Factors' statistical significance and their evolution overtime is of major importance since it can be an indicator that these factor-based trading strategies are already being exploited by other investors and that its alpha will soon disappear or that no longer exists, making it untrustworthy for future excess return generation. In addition, the disappearance of explanatory power of some factors means we can no longer understand what drives stock returns since incapability to explain stock returns makes it impossible for investors to estimate a reliable expected alpha, and therefore not allowing them to determine the best assets to invest.

This paper aims for the study of trading on well-known traditional factors by testing their statistical significance evolution over time and attempting to establish a valid reason for such results. For this study, we considered the purest of all factor investment definitions, a systematic approach to investments that continuously and substantially outperforms a benchmark. This paper needs to be consider of relevance since, for a technic that dates back to mid-70s, factor investing has now more than $900 billion invested in just about 820 ETFs, accordingly with FactSet. Furthermore, a survey by Invesco found that over 70 percent of institutional investors, surveyed in 2018, were using factor-based strategies and more than half were considering increasing its use.

This research paper is organized as the following: Section II analyses the core existent literature on the topics related to the model development and strategy definition. On Section III and IV is explained the data used and methodology defined, respectively. From Section V onwards, the results obtained are discussed, conclusions presented and laid some recommendations to investors that want to base their investment strategies on factor models.

2. Literature Review

2.1. Efficient Markets

Bachelier (1900), on his PhD thesis, modeled the first stochastic process in financial markets that, even though at the time was not considered of major importance it is today considered as such and also it is regarded as a pioneer advance at that time. This model had as fundamental principle that prices' speculation should be a fair game, which reflects on the expected profit of speculation being zero. Bachelier also included in his research the first description of the Random Walk Hypothesis and the principles of what today is titled as Efficient Market Hypothesis. His results pointed for the fact that current prices are a function of earlier fluctuations and current market position. Furthermore, he also affirmed that these previous fluctuations are determined by an infinite number of factors, consequently, creating the impossibility of modeling and forecasting exact prices. Recently, Cuthbertson and Nitzsche (2004) define a Random Walk as,

$$X_t = \delta + X_{t-1} + \varepsilon_{t+1}$$

where $\varepsilon_{t+1} \sim iid(0; \sigma_\varepsilon^2)$ and δ represents a drift, which is a weighted average of the probabilities of each price that the stock can assume in the next period.

Cowles (1933) analysis of the market as a whole concluded that it had outperformed its participants and affirmed that "The average forecasting agency fell [below] the average of all performances achievable by pure chance". Eleven years later, Cowles (1944) expanded his research, by using 6,904 expert forecasts over 15 years, determined once more that the results lacked ability to predict successfully the future course of the stock market.

Eugene Fama (1965) and Paul A. Samuelson (1965) developed a clearer and complete definition of the Efficient Market concept. This definition, with a broader acceptance of the concept, affirms that in an efficient market, where market prices incorporate all available information, these should be unpredictable. Fama (1970) maintained his focus on the Efficiency Market Hypothesis, but categorized empirical tests of efficiency into what he considered to be three different levels of market efficiency: i) Weak, ii) Semi-strong and iii) Strong. The weak form of market efficiency hypothesizes that current stock prices reflect all information that determined past prices and that it is impossible for investors to establish trading strategies based on past events. Furthermore, it also affirms that predicting future returns based on technical analysis is impossible since future stock prices follow a random walk. The semi-strong form of market efficiency is considered the most practical of the three forms of efficiency. It theorizes that current stock prices correct quickly to the release of all new public information and that neither technical nor fundamental analysis will allow an investor to outperform the market. Additionally, one can say that one of the few situations when an investor would be able to profit would be to use nonpublic information for trading purposes. The strong form of efficiency hypothesizes that both public and nonpublic information are already entirely reflected in the current stock prices, or if not, prices will quickly correct to their fair value. For example, if an investor holds information that is not available to the market, this person will trade on it until the trade is no longer profitable.

This hypothesis raised by Fama, has been since its formulation, a pillar of modern finance and up to this date continues to generate interest, as investors look for a better comprehension of the prices' generation process. Although the Efficient Market Hypothesis was consensus for many in the 70's, Lo and Mackinlay (1988, 1990) and others during the 80's, argued in favor of the rejection of the Efficient Market Hypothesis. To test their hypothesis, Lo and Mackinlay (1988, 1990) used variance ratio tests and analyzed US weekly market data from 1962 to 1985. These tests experienced on the fact that if prices follow a random walk, the variance of the logarithmic returns (logarithm of P_t minus logarithm of P_{t-1}, where P_t is the price at time t and P_{t-1} the price in the previous period) sampled with regular intervals of length time t and n times, has to be equal to the variance of the logarithmic returns sampled at intervals of length time t/n. Their conclusions pointed for the rejection of the Random Walk Hypothesis for the entire sample period and all sub-periods for a variety of aggregated returns indexes and size-sorted portfolios.

Even though currently considered outdated, it is important to mention the Random Walk Hypothesis (RWH) since the current Efficient Market Hypothesis (EMH) formulation relies upon some aspects of its existence. Even though the RWH differs from the EMH, mostly by the fact that the expected return of the RWH is zero while the EMH is conditional on a proper model for estimated returns (Fama, 1965), it is possible to establish a relation between both. Dupernex (2007) established the following condition to describe this: “If the Efficient Market Hypothesis holds, then future price movements will follow a purely random walk as new and unpredictable information emerges can describe the point of convection between both hypotheses”.

Recent research on the so-called Smart Beta strategies, which consists on investment strategies that apply a series of objective and rule-based screens (or factors) to the stocks in question in a search for an alpha, concluded on the possibility of these strategies also being negatively affected by its popularity (due to factor exploitation) as traditional factor investing is (Arnott, Beck, Kalesnik and West, 2016). Moreover, investors trying to attain the best performance possible can influence prices and make them reach higher values than what is their fair value, creating therefore a distorted reality when reviewing factors performance and statistical significance. Glushkov (2015) found no empirical evidence that, during the period of 2003-2014, Smart Beta funds outperformed their risk-adjusted benchmarks nor benefited from dynamic factor allocation.

2.2. Factor based investing

Subsequent to the work of Markowitz (1952), Treynor (1961, 1962) and Sharpe (1964), Fama and French (1992, 1996) built the prominent FF three-factor and FF five-factor model.

The Fama-French three-factor model is defined as,

$$r_{i,t} - rf_t = \alpha_{i,t} + \beta_1 * (rm_t - rf_t) + \beta_2 * SMB_t + \beta_3 * HML_t + \varepsilon_{it}$$

The Fama-French five-factor model is defined as,

$$r_{i,t} - rf_t = \alpha_{i,t} + \beta_1 * (rm_t - rf_t) + \beta_2 * SMB_t + \beta_3 * HML_t + \beta_4 * RMW_t + \beta_5 * CMA_t + \varepsilon_{it}$$

where $r_{i,t}$ is the total return of a stock or portfolio i at time t, rf_t is the risk free rate of return at time t, rm_t total market return at time t, $rm_t - rf_t$ is the excess return on the market, $r_{i,t} - rf_t$ is the expected excess return, SMB_t is the size premium, HML_t is the value premium, RMW_t profitability premium, CMA_t the investment factor that reflects difference between two portfolios defined as Conservative and two portfolios defined as Aggressive and ε_{it} the residual.

These models used to be able to explain stocks' returns in the past up to 90% (Yaşar, 2018) but its accuracy appears to be decreasing. For example, the size factor, SMB_t, which represents the premium existent between stocks with small and big market capitalization, appears to be disappearing. A reason for this might be the fact that investors have seen the available information for smaller companies increase, which gave them a wider range of companies to invest informed.

Harvey, Liu and Zhu (2016) determined that a large number of factors have explanatory power when it comes to justify abnormal stock returns, but further research by Freyberger, Neuhierl and Weber (2017) on factor investing proved that only a small number of factors can approximate stock returns in comparison to a large number of characteristics. Furthermore, Freyberger, Neuhierl and Weber found that out of twenty-four characteristics, only six to eleven provide independent information after resourcing to adaptive group least absolute shrinkage and selection operator.

Up to this moment, more than 300 factors have been identified as factors with value to explain stocks abnormal returns (Harvey and Liu, 2015). Although the number of factors identified is significant, the same cannot be said about their significance when tested statistically (Cazalet and Roncalli, 2014) (Hsu, Kalesnik and Viswanathan, 2015). When proven significant, and considered reliable for investment purposes, investors need to take into consideration that trading on this new factor will affect its returns, and if exploited on significantly large scale it can possibly be the origin of its disappearance (Ang, 2014). Ang (2014) recognized enormous losses for value, volatility and momentum factors during the last financial crisis. Additionally, also defended that factor investing is historically riskier in periods of poor market performance, with the factor premium acting as a long run compensation. Asness, Frazzini, Israel, Moskowitz and Pedersen (2018) found that only by controlling for the factor known as quality-minus-junk, a statistically significant premium exists for the size factor. Without this control, the statistical significance of the factor tends to vary mostly negatively overtime. For the period in question on our paper, this factor revealed to be statistically insignificant.

2.3. Long-short strategies

According with Shubert (2006), investors have been using long-short portfolios successfully for a long time. The idea behind a long-short portfolio consists in going long (buying) on stocks expected to outperform and short (selling) stocks expected to underperform. This strategy provides, theoretically, investors with a safeguard since these stocks will create a market neutral portfolio in relation to amount invested. Since long-short portfolios are likely to present a low or negative correlation with the market (Beaver, McNichols and Price, 2016), there is a significant improvement in performance when compared with a long only portfolio. Considering all this, it was decided that it would be appropriate to construct a long-short instead of a long-only portfolio.

3. Data

The data used to create this sample was all collected from Thomson Reuters Eikon, Bloomberg L.P. and Kenneth R. French Data Library. The data obtained from Thomson Reuters Eikon were the periods in which the stocks were components of the S&P100 Index (OEX) and the respective International Securities Identification Number (ISIN). The daily stock prices and indexes quotes required were obtained from Bloomberg L.P., as well as their respective ISIN. By combining both data sets, it was possible to create a unique dataset that ensured that only tradable stocks were considered for inclusion in the portfolio at any point in time and also ensure the exclusion of stocks not considered on the OEX index at the time of portfolio construction. Since certain information points were unavailable, due to extraordinary events (such as trading halts, M&A deals or exclusion from the index), it was assumed on these days stocks were trading at the previous day's closing price. An example of a well-known event that generated a situation like this was the death of six executives of Sundance Resources Ltd (an Australian company) after a plane crash. Expecting a stock price crash and unfair arbitrage by insiders, the company reached the Australian Securities Exchange and requested a trading halt until the information was widely spread to the public. The factors considered for this model were all retrieved from Kenneth R. French Data Library, except for the Sharpe factor.

4. Methodology

4.1. General definitions and conventions

The following indices will be used throughout this paper:

$t = rolling\ index\ that\ controls\ for\ the\ different\ days$;

$i = rolling\ index\ that\ controls\ for\ the\ individual\ stocks\ to\ be\ considered.$

The daily stock returns ($r_{i,t}$) were obtained as the natural logarithmic ratio between the closing price of the current day ($P_{i,t}$) and the closing price of the previous day ($P_{i,t-1}$), minus one. Mathematically it can be expressed as:

$$r_{i,t} = \ln\left(\frac{P_{i,t}}{P_{i.t-1}}\right) - 1$$

The market daily returns ($rm_{i,t}$) was computed with a similar methodology,

$$rm_{i,t} = \ln\left(\frac{Pm_{i,t}}{P_{i.t-1}}\right) - 1,$$

where $Pm_{i,t}$ is the current market index level and $Pm_{i,t-1}$ is the previous day's index level. During the following tests, the market index will be defined as the S&P100 (comprised by the 100 major blue chip companies in the US) or the S&P500 (comprised by the 500 biggest companies in the US).

The base considered to annualize the results presented was of two hundred and sixty days.

4.2. Factors Employed

The factors utilized were selected taking into consideration the importance that each can have individually, as well as collectively when predicting stock returns. One can claim that the factors retrieved for this paper were computed on a different set of stocks (it includes all NYSE, AMEX and NASDAQ stocks), which amounts to more than 3,000 stocks. Despite the fact that this is true, if factor trading is a reality, it shall be transversal to all markets and not exclusive to one.

The subsequent presented factors from the Kenneth R. French Data Library were chosen to allow the model to have the best chance of performing well, since it exposes it to different factor investing styles on well-known factors. Furthermore, it was also considered the fact that these factors have been, as previously presented, in the middle of strong divergences on their reliability for trading.

Unlike other factors, it was decided to include the Sharpe Ratio as an explanatory variable to represent market sentiment since it produces positive skewness and negative excess kurtosis for "well-behaved" investments and negative skewness and positive excess kurtosis for the bad ones (Prado, 2012), giving the model the best chance to perform at its best. Furthermore, it was decided a one hundred and thirty days window for this, incorporating market sentiment for the past six months.

The size factor, $SMB_{i,t}$ (Small Minus Big), consists on the average return on three small portfolios minus the average return on three big portfolios. Mathematically it can be defined as,

$$SMB_{i,t} = \frac{1}{3} * (Small\ Value + Small\ Neutral + Small\ Growth) - \frac{1}{3} * (Big\ Value + Big\ Neutral + Big\ Growth)$$

These six portfolios were formed daily by considering as Growth the bottom 30% companies of the book equity to market equity ratio, Neutral the ones between the 30% and 70% mark and as Value the remaining ones.

The value factor, $HML_{i,t}$ (High Minus Low), is the average return on two value portfolios minus the average return on two growth portfolios. It can be expressed as,

$$HML_{i,t} = \frac{1}{2} * (Small\ Value + Big\ Value) - \frac{1}{2} * (Small\ Growth + Big\ Growth)$$

These four portfolios were formed daily by considering as Growth the bottom 30% companies of the book equity to market equity ratio, Neutral the ones between the 30% and 70% mark and as Value the remaining ones.

The market factor, $MKTRF_{i,t}$, is the market excess return weighted on a market capitalization basis of all United States incorporated companies on the Center for Research in Security Prices (CRSP) and listed on the NYSE, AMEX or NASDAQ subtracted from the one month interest rate on a United States Government Treasury bill.

The momentum factor, $MOM_{i,t}$, is the average return on two high prior return portfolios minus the average return on two low prior return portfolios, i.e.,

$$MOM_{i,t} = \frac{1}{2} * (Small\ High + Big\ High) - \frac{1}{2} * (Small\ Low + Big\ Low)$$

These four portfolios were formed daily, considering prior return, using the percentiles 30^{th} and 70^{th} of the previous NYSE components daily returns.

All previously presented portfolios for the size, value and momentum factors as “Big” and “Small” were defined taking as separation point the median market equity.

The market performance factor, $Sharpe_{i,t}$, was defined as the one hundred and thirty days rolling Sharpe ratio for the S&P100, to allow the incorporation of market performance for the past six months.. Analytically it can be defined as,

$$Sharpe_{i,t} = \frac{rm_{130,i,t}}{\sigma_m},$$

where σ_m is the market one hundred and thirty days rolling volatility and $rm_{100,i,t}$ the return during the previous one hundred and thirty days. The rolling volatility can be defined as,

$$\sigma_m = \sqrt{\frac{\sum_{i=1}^{n}(r_{i,t} - r_{130})}{n-1}},$$

where n is one hundred and thirty and r_{130} an average return of the prior one hundred and thirty days.

4.3. Model

4.3.1 Model development

After retrieving the required trading factors data from the Kenneth R. French Online Data Library (and creating a lag of one day to account for forward looking bias), prices from Bloomberg LP and defining the periods when it was possible to trade stocks (resourcing to Thomson Reuters Eikon), it was possible to compute daily returns. Following this, resorting to Microsoft Excel function Linest, it was possible to compute the $\beta_{Factor,i,t}$ (beta) for each factor, using a one hundred and thirty days rolling window. The one hundred and thirty days window was chosen so that the model would be influenced by taking into consideration the stock price evolution and factors over the past six months. The Linest function from Microsoft Excel computes statistic parameters, using the linear least squares method, for a multiple linear regression that creates the best fit for the dataset in question. The multiple linear model in question can be defined as,

$$r_{i,t} = \alpha_{i,t} + \beta_{MOM,i,t} * MOM_{i,t} + \beta_{SMB,i,t} * SMB_{i,t} + \beta_{HML,i,t} * HML_{i,t} + \beta_{\mathrm{MKTRF},i,t} * \mathrm{MKTRF}_{i,t} + \beta_{Sharpe,i,t} * Sharpe_{i,t} + \varepsilon_{i,t},$$

where $\varepsilon_{i,t}$ represents the residual value for stock i at time t and follows a normal distribution, $\beta_{Factor,i,t}$ represents the Beta coefficient for the factor for the stock i at time t and $r_{i,t}$ the expected return for stock i at time t. The least square estimators of the coefficients in the linear prediction function are then obtained by minimizing the sum of the difference of the squared deviations between the model's expected values and the observed, i.e.,

$$min \sum_{i=1}^{N}((\alpha_{i,t} + \beta_{MOM,i,t} * MOM_{i,t} + \beta_{SMB,i,t} * SMB_{i,t} + \beta_{HML,i,t} * HML_{i,t} + \beta_{\mathrm{MKTRF},i,t} * \mathrm{MKTRF}_{i,t} + \beta_{Sharpe,i,t} * Sharpe_{i,t}) - r_{i,t})^2.$$

The stock's daily residual value, i.e., the potential that the stock is considered to have for a relative gain, was defined as the difference between the model's expected returns for each stock on a daily basis and the actual daily returns. We will consider this factor to be the potential return a stock has, given our predefined model.

$$\varepsilon_{i,t} = (\alpha_{i,t} + \beta_{MOM,i,t} * MOM_{i,t} + \beta_{SMB,i,t} * SMB_{i,t} + \beta_{HML,i,t} * HML_{i,t} + \beta_{\mathrm{MKTRF},i,t} * \mathrm{MKTRF}_{i,t} + \beta_{Sharpe,i,t} * Sharpe_{i,t}) - r_{i,t}.$$

Technics similar to this have become very popular in machine learning (Yang, Liu and Wu, 2019) and algorithmic trading (Park and Phadnis, 2017). These models, just as the one presented on this paper, try to capture value that is not yet reflected in the stocks' prices, by establishing a connection between an (or several) underlying indicator(s).

For stock picking purposes, it was decided not to rank stocks based on their daily residual values but on the sum of the last five daily residuals to create a more stable portfolio (a long-short portfolio that is composed by 10 long stocks and 10 short stocks, created with the intention of giving the previously presented hypothesis the chance of failure), when it comes to its components and therefore reducing transaction costs. The five days rolling window was decided based on the usual number of days that exist in a week. This can mathematically be expressed as,

$$Sum\_\varepsilon_{i,t} = \sum_{i=1}^{5} \varepsilon_{i,t} = \varepsilon_{i,t} + \varepsilon_{i,t-1} + \varepsilon_{i,t-2} + \varepsilon_{i,t-3} + \varepsilon_{i,t-4}.$$

The ranking system was designed to provide a low score to stocks that had a superior value on the $Sum\_\varepsilon_{i,t}$., i.e. the ranking score was defined in an ascending order. One was attributed to the stock with the best chance of registering a relative increase in value (the stock with the highest positive residual) and higher values were attributed to stocks with the lowest potential for a relative gain. It is important to remember that on this model top performers, i.e. the stocks with the biggest distance to expected return the model predicted, were bought (long), while the worst performers, i.e. the stocks with the lower residuals, were sold (short). It was considered the difference of one day between the computation of the ranks and the execution of the buy or sell orders, to ensure that we do not incur in forward looking bias and, therefore safeguarding against overly positive results.

The portfolio's daily returns ($r_{portfolio,t}$) are computed as an equally weighted average of daily realized return of each individual stock that is a portfolio component. This can be synthetized as,

$$r_{\text{portfolio,t}} = \frac{1}{n} * \sum_{i=1}^{n} r_{i,t}\ , where\ n\ is\ the\ number\ of\ stocks.$$

This portfolio composed by 10 long stocks and 10 short stocks, so that it comprises the first and last 10% percentile, has as objective to provide the previously presented hypothesis with the best chance of failure since it takes into consideration all factors at once for stock picking purposes.

4.3.2. Transaction costs

It is important to consider transaction costs during investment strategies' development, since it will reduce individual stocks' and portfolio's daily returns, consequently decreasing total returns. This profitability reduction can be of such magnitude that it can turn extremely good trading strategies into investment strategies that are not reliable to be executed (Korajczyk and Sadka 2002).

In this paper, it was considered that transaction costs were fixed at 0.015% (Engle, Ferstenberg and Russell, 2008) per trade per security (buying and selling costs are assumed to be identical). This value, besides accounting for the transaction costs, it accounts for the bid ask spread. This important spread has not yet been reflected prior to this, since the returns were computed resorting to daily closing prices and not bid and ask prices.

5. Results and analysis

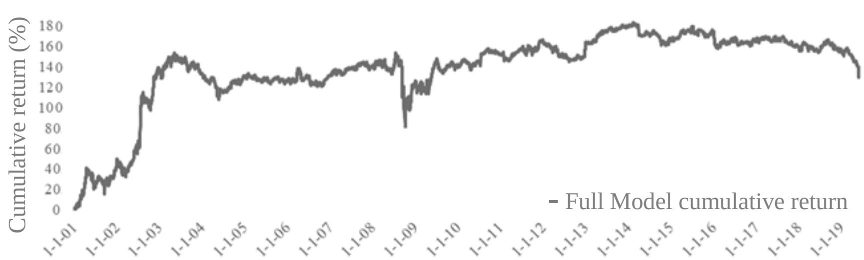
This section is dedicated to the analysis and discussion of the results obtained after the development of the earlier presented model and methodology, and further individual factor analysis. At this point is imperative to reiterate that our research question is if factors can present consistent returns overtime and analyze their statistical evidence. For such, several portfolios were constructed with the objective of assessing this. If our strategy results match the anticipated, we expect strong financial results at the start of the sample that decrease overtime and low statistical evidence.

The investment period was established to have a start on January 1st 2001 and end on May 31st 2019, which implies an analysis period of 18 years and 5 months. The starting point for this sample was selected taking into consideration the existence of a market that one can say to be bullish most of the time, which is favorable for factor investing (Ang, 2004), but still captures part of the dot-com bubble (Tech bubble), the 21st century financial crises, the impact of terrorism attacks in American soil and the switches of political parties in the White House. The ending point was determined as the last complete month at the start time of the writing of this paper.

5.1. Model Results

Trading on the previously defined strategy, which consists on trading on the potential value that stocks have for relative price movements accordingly with the model, with the 10 stocks with the best potential for appreciation longed and the 10 stocks more probable to lose value shorted, pointed for a total return of 138.86%, which represents an annualized return of 7.11% and annualized standard deviation of 20.32%, resulting on an Info Sharpe ratio of 0.35. These results are relevant for our study since they provide the best premise for the factors to perform well, and therefore prove our hypothesis wrong. From now on, this portfolio will be designated as "Full Model" and its cumulative return of can be perceived on the graphic below.

Graph 1 – Full Model cumulative return



As it can be observed from the graph above, a significant part of the return obtained by the model results from the three first years in question. This might be an indicator that these factors have been losing predictive power, which means lower returns as the years pass. Furthermore, another relevant aspect from visual analysis of the graph above is the drop during the year of 2008, when our model loses almost 60 p.p., while the S&P100 loss is only around 37 percent. This might be an indicator that the model has stopped working around this period since during the years of

2001 and 2002, the index lost more than 38 p.p. but our model was able to return almost 100 p.p. Furthermore, Ang (2004) defended that factor investing is historically risker in periods of poor market performance.

To better understand the results obtained by the model, it was firstly performed an overall significance test. The selected statistical test was the F test since it provides a better understanding if the factors considered in the model explain the results obtained by testing their overall significance. Considering the previously model defined as,

$$r_{i,t} = \alpha_{i,t} + \beta_{MOM,i,t} * MOM_{i,t} + \beta_{SMB,i,t} * SMB_{i,t} + \beta_{HML,i,t} * HML_{i,t} + \beta_{MKTRF,i,t} * MKTRF_{i,t} + \beta_{Sharpe,i,t} * Sharpe_{i,t} + \varepsilon_{i,t}$$

we would be testing,

$$H_0 : \beta_{MOM,i,t} = \beta_{SMB,i,t} = \beta_{HML,i,t} = \beta_{MKTRF,i,t} = \beta_{Sharpe,i,t} = 0$$

$$H_1 : \exists \beta_{j_{i,t}} \neq 0, \qquad where\ j = MOM, SMB, HML, MKTRF, Sharpe$$

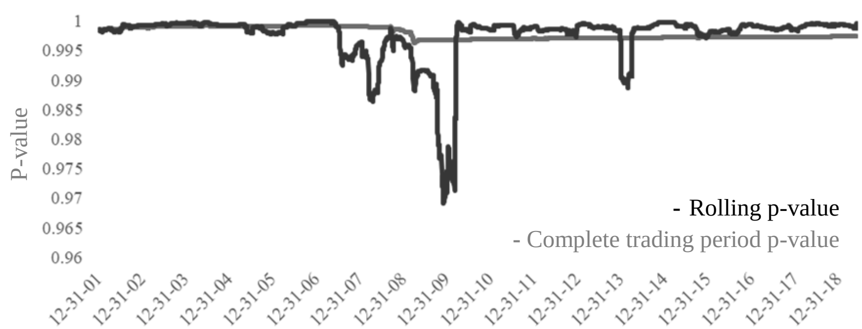
In other words, we would be testing if we reject or not that all coefficients associated to the explanatory variables are at the same time equal to zero.

The previously mentioned F statistic can be defined as $F = \frac{\frac{R^2}{k-1}}{\frac{1-R^2}{n-k}} \sim F_{(n-1,n-k)}$, where k is the number of tested coefficients and n the number of observations.

The analysis of the 260 days rolling alpha (with alpha being defined as the level of significance) at 0.05 (or 5%) significance level pointed for the failure of the overall significance of the variables (i.e., we fail to reject H_0) 100% of the times. Here, we can observe strong evidences that this cluster of variables has small to no explanatory power during most of the trading period.

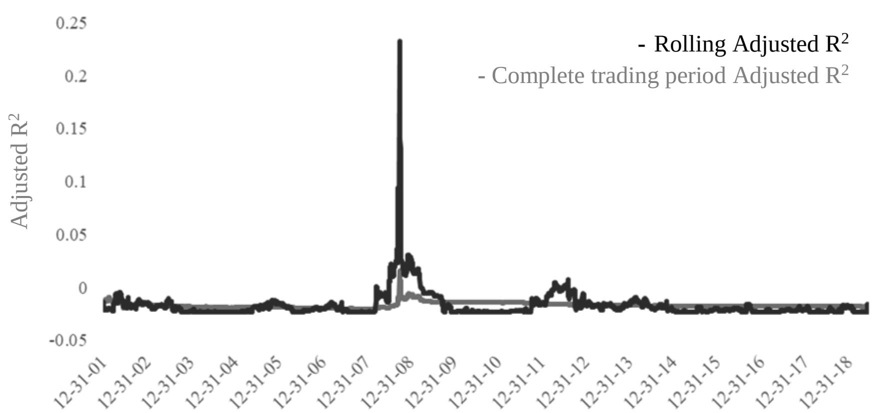
Further analysis of the p-value, considering it now for the whole sample period, pointed for similar, but even stronger (constant) results. On the graphic below, it is possible to observe the evolution of both the rolling p-value and the p-value for the complete trading period.

Graph 2 – Full Model p-value



A measurement of fit of the model to the data set is the $Adjusted\ R^2$, a modified version of the R^2. This measure is a tool to better understand the explanatory power of models that contain different numbers of predictors. Just like the p-value, the $Adjusted\ R^2$ was tested on a fixed and rolling basis, represented in the graphic below.

Graph 3 – Full Model $Adjusted\ R^2$



As it can be observed, the rolling $Adjusted\ R^2$ demonstrates a lower value than what it would be desirable to establish a reliable trading strategy. Furthermore, even though it appears that the model has no explanatory capabilities for the results obtained, it is important to highlight that the $Adjusted\ R^2$ value registers significant increases just prior to periods when the markets can be said to be bearish. A possible reason for this might be a shift in investors' investment styles to what can be said to be more traditional styles in times of uncertainty. Additionally, it can be observed that the fixed $Adjusted\ R^2$ remains below zero during most of the period in analysis, with a slight but irrelevant upward trend (it is said to be irrelevant since this trend is strongly influenced by the 2007-2009 period).

5.2. Factors Analysis

Taking into consideration the results earlier obtained, the next logical test would be a deeper factors' analysis. This analysis was executed by performing statistical test on the individual significance of each explanatory variable individually to verify if any statistical evidence for the registered returns exists, which are expected to fluctuate over the period in question, and provide a better understanding of the Full Model results. The individual significance tests were performed using trading models similar to the following:

$$r_{i,t} = \alpha_{i,t} + \beta_j * j_{i,t} + \varepsilon_{i,t}, where\ j = MOM, SMB, HML, MKTRF, Sharpe.$$

The selected test for this was the t test since it is the most approprieted for the intended. These tests consisted on:

$$H_0 : \beta_{j_{i,t}} = 0 \text{ and } H_1 : \beta_{j_{i,t}} \neq 0, where\ j = MOM, SMB, HML, MKTRF, Sharpe.$$

The subsequent table contains a summary of the previously described statistical tests and its financial results.

Table 1 - Summary of results for models with transaction costs

Models	Significant at α =5%?	Cumulative Return	Annualized Return	Standard Deviation	Info Sharpe Ratio
SMB	No	-62.05%	-3.36%	24.95%	-0.13
HML	No	84.46%	4.57%	24.97%	0.18
MKTRF	No	14.74%	0.80%	24.20%	0.03
MOM	No	25.23%	1.36%	24.18%	0.06
Sharpe	No	-22.65%	-1.23%	24.00%	-0.05

The single factor models all revealed to have no significance at a significance level of 5% for the factors but the results seem interesting and revealing. The factor that represents companies' market capitalization (or size), SMB, was the factor that registered the worst of all results, with an Info Sharpe ratio below -0.1, which might be an indicator that the factor no longer presents a reliable premium. Just like the SMB, the Sharpe factor also registered a negative Info Sharpe ratio, however less than half the value registered be the previous factor. The factors MKTRF and MOM both have an Info Sharpe within a range with a length inferior to 0.03 and similar volatilities, which can possibly be an indicator of both factors been regularly traded in the market as a group (these factors' daily returns present a correlation close to 0.95 for the complete period in question). Lastly, the best performing factor was value, HML, returning the relatively low but comparably high (to the other models) Info Sharpe ratio of 0.18, showing potential to be analyzed when trading costs are excluded.

It is important to notice the following: i) several factors p-values appear to move prior to major market events (possibly due to investors shifting to strategies perceived as more secure) (Bordley R. and Luisa Tibiletti, 2015), ii) the factors MOM and MKTRF appear to be similar when it comes to the rolling p-values' evolution (recording a strong correlation, superior to 0.76 for the whole period), iii) all single factor models' $Adjusted\ R^2$ have strong and consistent movements prior, during or after moments of major market activity (mostly in the case of bearish markets and possibly for the same reasons presented above for the p-values movements) (Bordley R. and Luisa Tibiletti, 2015), and iv) the factors' financial results are not consistent overtime, with individuals' cumulative returns depending strongly on the first two to three years of the sample (see appendix).

5.3. Trading excluding transaction costs

It was considered that the next appropriated step to simulate, after the results obtained and analysis, would be to exclude transaction costs for a fair comparison with the S&P100 and S&P500, since both do not have such rebalancing considerations on their performance. A summary of the results for single factors and multi-factor models are presented on the table below. These results were computed taking into consideration an equal trading period to the one defined for the model with transaction costs, so that later it is possible to compare both.

Table 2 - Summary of results for models without transaction costs

Models	Significant at α =5%?	Annualized Return	Standard Deviation	Info Sharpe Ratio
S&P500	-	4.07%	18.81%	0.22
S&P100	-	3.17%	18.68%	0.17
Full Model	No	10.87%	20.31%	0.53
MOM	No	7.00%	24.18%	0.29
MKTRF	No	6.43%	24.20%	0.27
HML	No	6.44%	24.97%	0.26
Sharpe	No	4.39%	24.00%	0.18
SMB	No	2.28%	24.94%	0.09

As it could be expected, the results for all models have improved to the point of four of the six models have now a better performance than the S&P500 and five a better performance than both the S&P500 and S&P100. The Full Model is now outperforming the S&P500 by more than 6.5 p.p. per year, with a similar volatility level to the large caps index. Furthermore, its Info Sharpe ratio is now above 0.53, a value that one can say to be somewhat reliable to execute trades on. All this shows that the stocks picked by the model are good but trading costs erase trading opportunities.

When it comes to the single factor models, the MOM, MKTRF and HML registered higher returns but also higher volatility, which returned an Info Sharpe ratio similar to the S&P500's. As it can be observed from the graphics on the appendix, the individual factors still do not demonstrate any capabilities to explain the results obtained.

When it comes to statistical evidence, both the $Adjusted\ R^2$ and p-value registered for all models similar results to the ones for the models with transaction costs (see appendix). Therefore, we can say that transaction costs are not having a relevant impact on our statistical results.

An indicator of trading frequency is the difference between the annualized return of the model without trading costs and the one with trading costs. These differences can be found on Table 3.

Table 3 – Analysis of the impact of transaction costs

Models	Annualized Return without trading costs	Annualized Return with trading costs	Difference
Full Model	10.87%	7.11%	+ 3.76 p.p.
MOM	7.00%	1.36%	+ 5.64 p.p.
MKTRF	6.43%	0.80%	+ 5.63 p.p.
HML	6.44%	4.57%	+ 1.87 p.p.
Sharpe	4.39%	-1.23%	+ 5.46 p.p.
SMB	2.28%	-3.36%	+ 5.64 p.p.

Intuitively one can say that the models that have the highest differences are the ones with the highest trading frequency, i.e., the ones that change its components more often. The portfolio with the lowest trading frequency is the value one, HML, with the difference below 2 p.p., while all others single factor models have differences of roughly 5.5 p.p.. As expected, the Full Model ranges somewhere in the middle of the single factors' results with a difference of 3.76 p.p., which reflects on lower trading frequency than all portfolios except the HML.

It was important to perform the tests above since it was hypothesized that transaction costs could be influencing factors' performance and their statistical evidence. This revealed to be truth for the first but not for the second hypothesis.

6. Conclusions and Recommendations

In this paper, the factors size (SMB), growth (HML), market (MKTRF), momentum (MOM) and market's info Sharpe ratio (Sharpe) were tested taking into consideration different contexts (using multi and single factor models) and assumptions (including and excluding transaction costs), to analyze factor performance and statistical significance overtime. The objective of these tests was to demonstrate that factor is currently not a reliable strategy under realistic assumptions. The Full Model, the one that includes all five factors, showed relatively strong financial results in comparison with both considered benchmarks (S&P100 and S&P500), but still no statistical evidence was found that the these explanatory variables can explain the results obtained and produce future similar results. As it can be noticed from visual observation of the cumulative return's graph, a large portion of the returns results from a short period of the full sample, i.e. it is not a sustained result overtime, but minor periods that support these results. This, as previously discussed and concluded by others (McLean and Pontiff, 2016), might be caused by investors learning from published academic research and trying to arbitrage on this knowledge from that point on. To highlight, remains the fact that we are in the presence of a trading strategy that would only be financially reliable for this period under the unrealistic assumption that it could be executed without any costs, without any statistical evidence being found that could support these results.

When it came to the individual performance and statistical significance analysis, these tests were executed to provide a better understanding of what was contributing for the results of the Full Model. As expected, the performance of these models was similar to the Full Model, being strongly dependent from the first half of the sample. Unlike anticipated, the statistical significance did not decrease overtime but was persistently low during all the period in analysis except during major bearish market movements. The size factor was constantly the one that revealed to perform the worst, with and without transaction costs. The HML, MKTRF and MOM factors showed to have similar results for the return and standard deviation for the model without transaction costs. The same cannot be said for the models with transaction costs, where the HML revealed to be a low frequency rebalancing portfolio comparatively to all others.

The main objective of this paper was to analyze the statistical significance of well-known factors. As previously concluded, these factors did not reveal any signs of being statistically significant when tested individually or as part of a complete model but were capable of generating interesting financial results. These results are similar to the ones obtained by Harvey and Liu (2015). This creates evidences that these factors no longer allow investors the opportunity of collecting on a premium, possibly because they are/were traded on with such intensity that this premium was reduced to a residual value that no longer justifies trading activity. Some reasons for this increase in trading volume are perhaps the recent upsurge of high frequency trading, which facilitated traders with an easier option to explore factors, the creation of dark pools over the past decade, or even the exploration of the so-called Smart Beta using ETF passive strategies. All these affect positively market liquidity, therefore reducing premiums and removing opportunities for the average trader.

Throughout the tests, transaction costs had a significant impact on the final performance of the Full Model strategy, demonstrating this that the portfolio in question is a relatively high-rebalancing frequency portfolio. To have results that are more accurate to real life trading activity, other costs should also be taken into consideration, such as financing, something that was ignored in this report. The presence of such costs should be expected to have an impact with similar magnitude to the one caused by the inclusion of trading costs in the models without trading cost. Another point of improvement for this model is the prices used for the computation of the returns. During this paper, were used the daily close prices and the bid-ask spread was compensated with the use of transaction costs slightly higher than the actually required, when in fact there was no need for such if the bid and ask prices were used straight away. Lastly, the number of companies used for these tests can be said to be small since it only included the components of the S&P100. This shall be considered of relevance since, besides being a relatively small number of companies, we are only looking at companies that have massive market capitalizations.

It would be interesting to test the previous statistics but taking into consideration the existence of industries that constitute the indexes, and test for each the significance of the factors. Furthermore, as it was previously raised the idea after visual analysis of the results, it would also be interesting to test the same factors and models but analyzing their capability to predict major financial movements, such as financial crisis.

References

- Ang, Andrew. 2014. *Asset management: A systematic approach to factor investing*. New York: Oxford University Press.
- Arnott, Robert D., Noah Beck, Vitali Kalesnik, and John West. 2016. "How Can 'Smart Beta' Go Horribly Wrong?" Working paper available at SSRN.
- Asness, Clifford S., Andrea Frazzini, Ronen Israel, Tobias J. Moskowitz, and Lasse H. Pedersen. 2018. "Size Matters, If You Control Your Junk." *Journal of Financial Economics,* 129 (3): 479-509.
- Bachelier, Louis. 1900. *"Théorie de la speculation." Annales scientifiques de l'École Normale Supérieure,* 17: 21-86.
- Beaver, William, Maureen McNichols and Richard Price. 2016. "The costs and benefits of long-short investing: A perspective on the market efficiency literature." *Journal of Accounting Literature,* 37: 1–18.
- Bordley, Robert, and Luisa Tibiletti. 2015. "hy Investors are Loss Averters During Bull Markets and Gain Seekers During Bear Markets?" Working paper available at SSRN.
- Cazalet, Zélia and Thierry Roncalli. 2014. "Facts and Fantasies about Factor Investing." Working paper available at SSRN.
- Cowles, Alfred. 1933. "Can Stock Market Forecasters Forecast?" *Econometrica: Journal of the Econometric Society,* 1: 309-324.
- Cowles, Alfred. 1944. "Stock market forecasting." *Econometrica: Journal of the Econometric Society,* 12: 206-214.
- Cuthbertson, Keith, and Dirk Nitzsche. 2004. *"Quantitative Financial Economics - Stocks, Bonds and Foreign Exchange."* New Jersey: John Wiley and Sons Ltd.

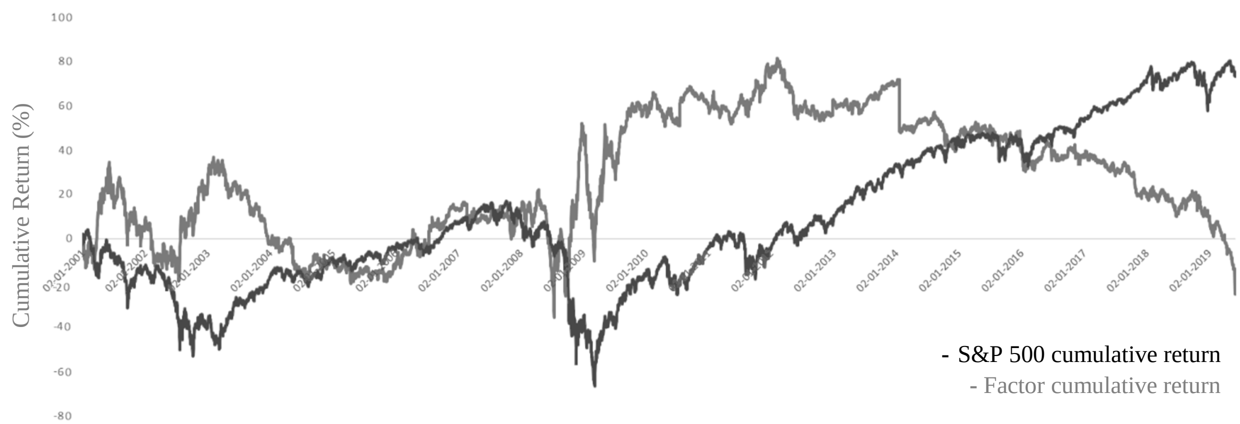
- Dupernex, Samuel. 2007. “Why might share prices follow a random walk?” *Student Economic Review*, 21.
- Engle, Robert F., Robert Ferstenberg and Jeffrey R. Russell. 2006. “Measuring and Modeling Execution Cost and Risk.” *The Journal of Portfolio Management*, 38 (2).
- Erdinç, Yaşar. 2018. “Comparison of CAPM, Three-Factor Fama-French Model and Five-Factor Fama-French Model for the Turkish Stock Market.” In *Financial Management from an Emerging Market Perspective*, ed. Guray Kucukkocaoglu and Soner Gokten, 69-93. Rijeka: InTech.
- Freeman, Richard B. 1993. “How Much Has Deunionization Contributed to the Rise in Male Earnings Equality?” In Uneven Tide: Rising Income Inequality in America, ed. Sheldon Danzinger and Peter Gottschalk, 133–63. New York: Russell Sage Foundation.
- Fama, Eugene F. 1965. “The Behavior of Stock-Market Prices.” *The Journal of Business*, 38 (1): 34-105.
- Fama, Eugene F. 1970. “Efficient Capital Markets: A Review of Theory and Empirical Work.” *The Journal of Finance,* 25 (2): 383-417.
- Fama, Eugene F., and Kenneth R. French, 2015. “Incremental variables and the investment opportunity set.” *Journal of Financial Economics,* 117(3): 470-488.
- Fama, Eugene F., and Kenneth R. French. 1992. “The cross-section of expected stock returns.” *The Journal of Finance,* 47: 427–465.
- Fama, Eugene F., and Kenneth R. French. 1996. “Multifactor Explanations of Asset Pricing Anomalies.” *The Journal of Finance*, 51(1): 55–84.
- Freyberger, Joachim, Andy Neuhierl and Michael Weber. 2017. “Dissecting Characteristics Nonparametrically.” *CESifo Working Paper Series*, 6391.

- Glushkov, Denys. 2015. “How Smart are 'Smart Beta' ETFs? Analysis of Relative Performance and Factor Exposure.” Working paper available at SSRN.
- Harvey, Campbell R. and Yan Liu. 2014. “Multiple testing in economics.” Working paper available at SSRN.
- Harvey, Campbell R., and Yan Liu. 2015. “Lucky factors.” Working paper available at SSRN
- Harvey, Campbell R., Yan Liu and Caroline Zhu, 2016. “... and the Cross-section of Expected Returns”, *Review of Financial Studies*.
- Hsu, Jason C. and Vitali Kalesnik. 2014. “Finding Smart Beta in the Factor Zoo.” Working paper available at Research Affiliates.
- Hsu, Jason C., Vitali Kalesnik and Vivek Viswanathan. 2015. “A framework for assessing factors and implementing smart betas strategies.” *The Journal of Index Investing*, 6(1): 89-97
- Hunstad, M. and Jordan Dekhayser. 2015. “Evaluating the Efficiency of 'Smart Beta' Indexes.” *The Journal of Index Investing*, 6(1): 111-121.
- Kenneth R. French - Data Library. 2019. “Current Research Returns”. Retrieved from Kenneth R. French Data Library http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/index.html.
- Korajczyk Robert A., and Ronnie Sadka. 2002. “Are Momentum Profits Robust to Trading Costs?” Working paper available at SSRN.
- Lo, Andrew W., and MacKinlay A. Craig. 1988. “Stock Market Prices do not Follow Random Walks: Evidence from a Simple Specification Test.” *The Review of Financial Studies,* 1: 41-66.
- Lo, Andrew W., and MacKinlay A. Craig. 1990. “When Are Contrarian Profits Due to Stock Market Overreaction?” *The Review of Financial Studies,* 3: 175–205.
- Markowitz, Harry. 1952. “Portfolio Selection.” *The Journal of Finance,* 7: 77-91.

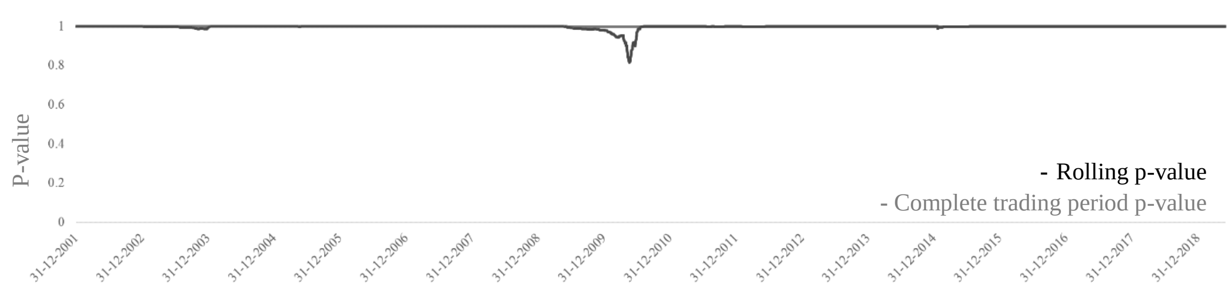
- McLean, R. David, and Jeffrey Pontiff. 2016. "Does Academic Research Destroy Stock Return Predictability?" *The Journal of Finance*, 71 (1): 5-32.
- Nayema K., Achola, and Peter Muriu. 2016. "Testing the three factor model of Fama and French: evidence from an emerging market." *European Scientific Journal*, 12(4): 211-236.
- Park, Sanghyun, and Kapil Phadnis. 2017. "Algorithmic Trading of Portfolios." Working paper available at SSRN.
- Prado, Marcos L. 2012. "The Sharp Razor: Performance Evaluation with Non-Normal Returns." Working paper available at SSRN.
- Quance, Stephen, and Blaise Warren. 2019. "Foundational concepts for understanding factor investing." Retrieved from https://apinstitutional.invesco.com/home/our-views-factor/foundational-concepts-for-understanding-factor-investing.
- Samuelson, Paul A. 1965. "Rational Theory of Warrant Pricing." *Industrial Management Review*, 6 (2): 13-39.
- Sharpe, William F. 1964. "Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk." *Journal of Finance*, 19: 425-442.
- Shubert, Leo. 2006. "Long-Short Portfolio." *Economic Analysis Working Papers*, 5(14).
- Treynor, Jack L. 1961. "Market Value, Time, and Risk." Unpublished manuscript dated 08/08/1961.
- Treynor, Jack L. 1962. "Toward a Theory of Market Value of Risky Assets." Unpublished manuscript dated fall 1962. A final version was published by Robert Korajczyk. 1999. "Asset Pricing and Portfolio Performance", *Risk Books*, 1: 15-22.
- Yang, Hongyang, Xiao-Yang Liu and Qingwei Wu. 2019. "A Practical Machine Learning Approach for Dynamic Stock Recommendation." Working paper available at SSRN.

Appendix

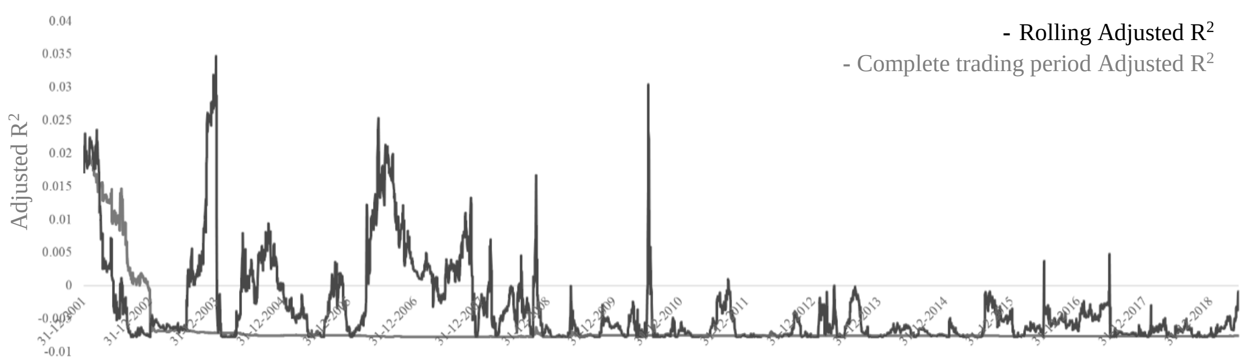
Appendix A – Cumulative Return for the Factor Sharpe with transaction costs



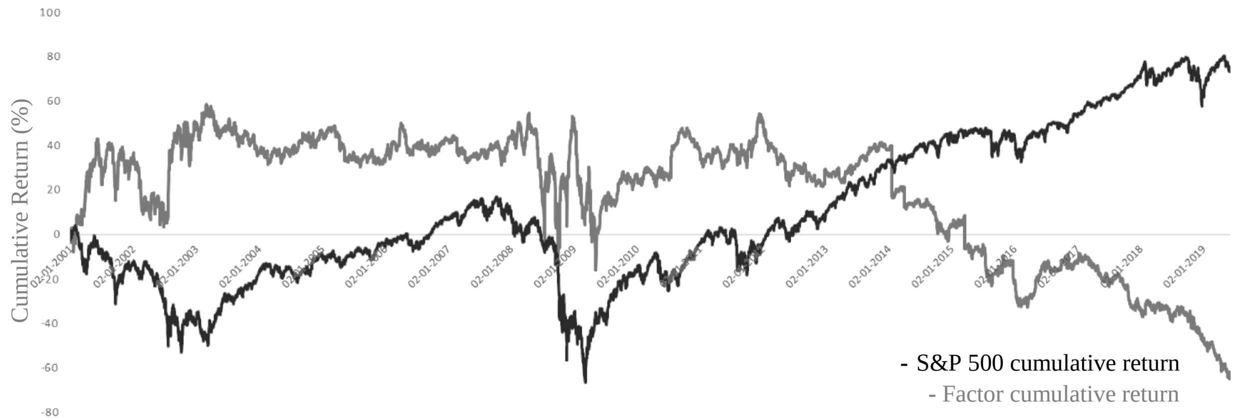
Appendix B – P-value of the Factor Sharpe with transaction costs



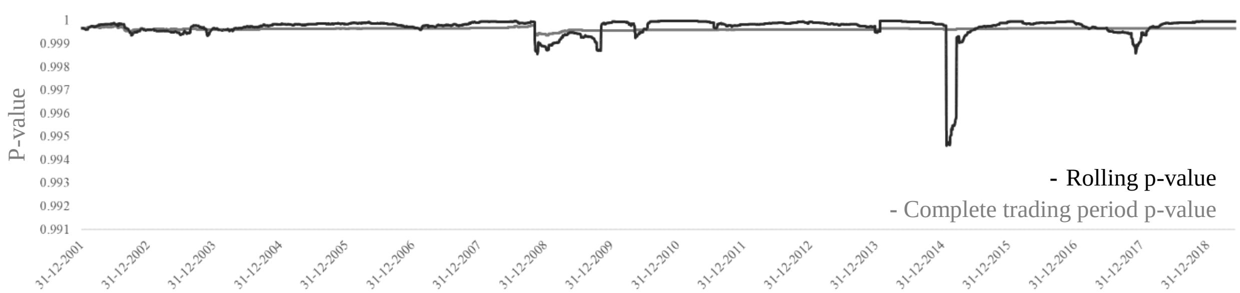
Appendix C – Adjusted R Squared of the Factor Sharpe with transaction costs



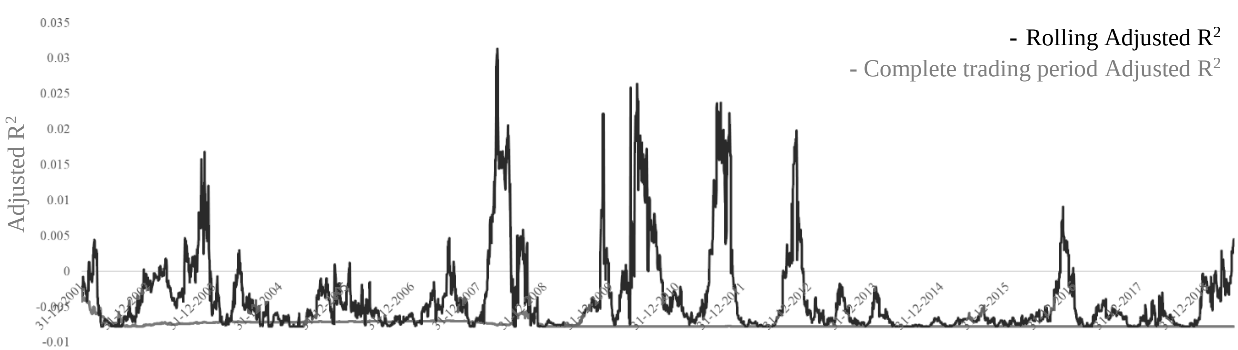
Appendix D – Cumulative Return for the Factor SMB with transaction costs



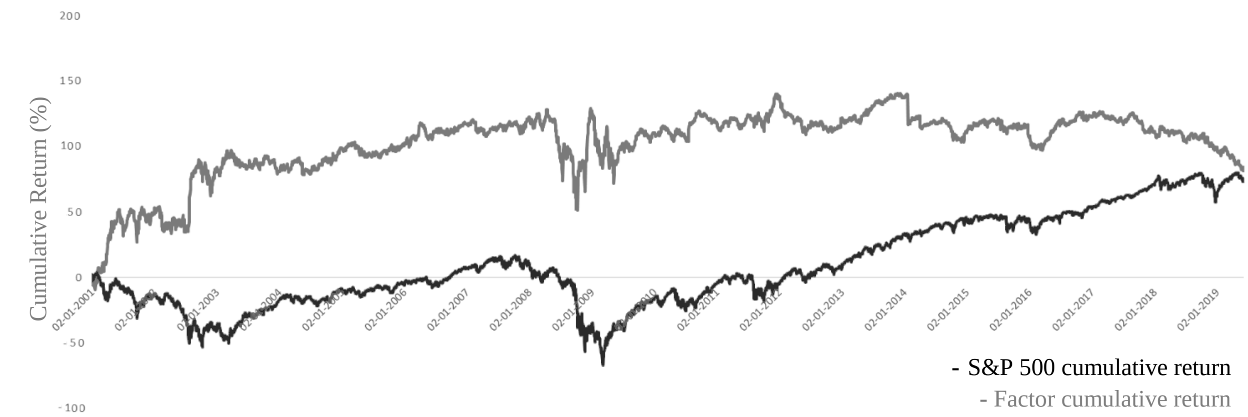
Appendix E – P-value of the Factor SMB with transaction costs



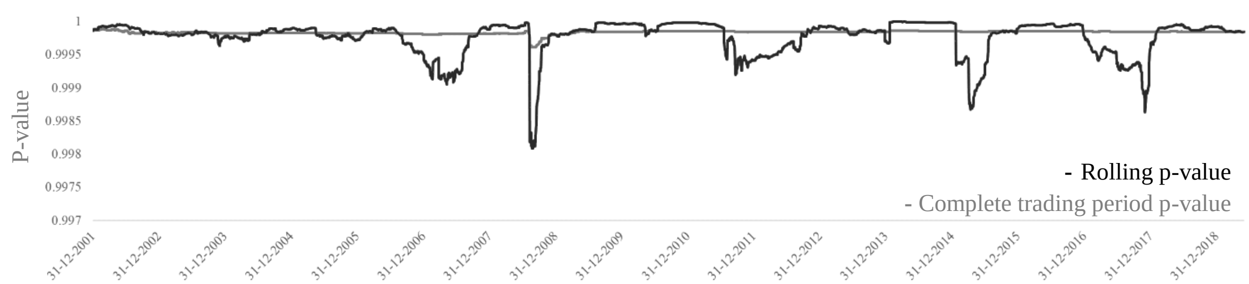
Appendix F – Adjusted R Squared of the Factor SMB with transaction costs



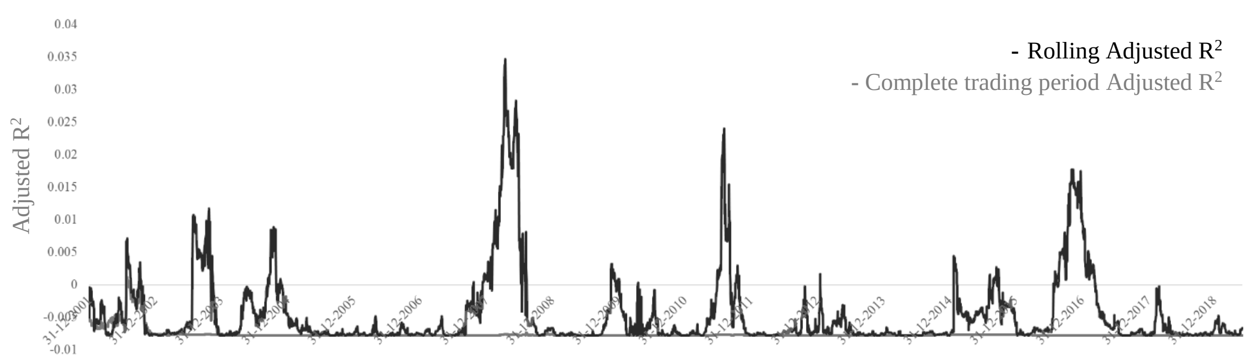
Appendix G – Cumulative Return for the Factor HML with transaction costs



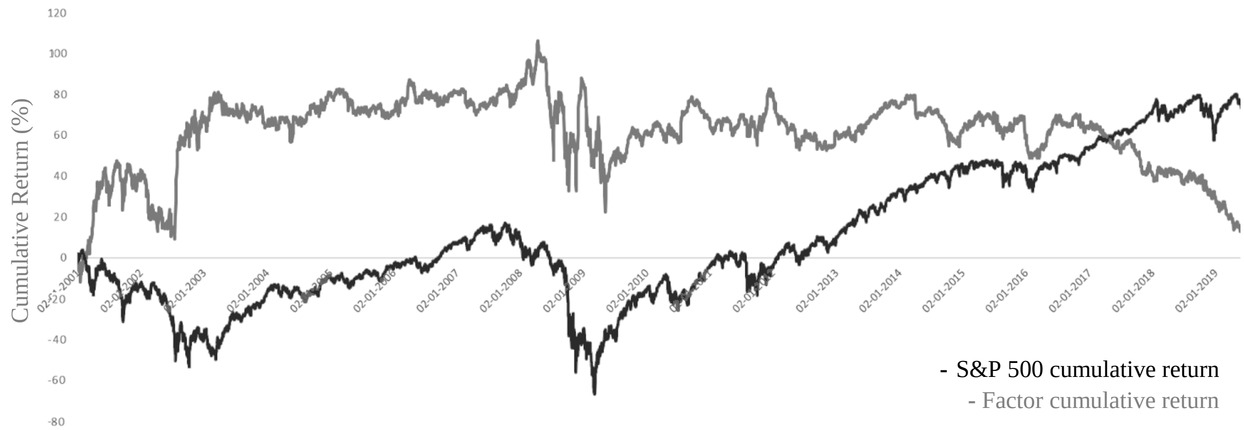
Appendix H – P-value of the Factor HML with transaction costs



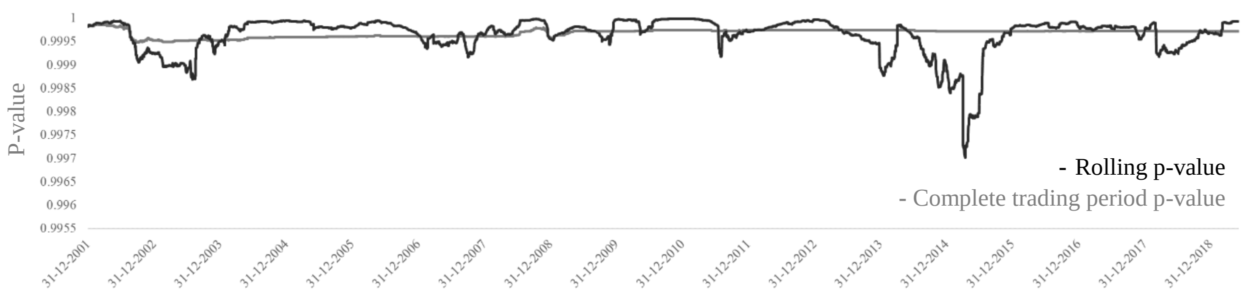
Appendix I – Adjusted R Squared of the Factor HML with transaction costs



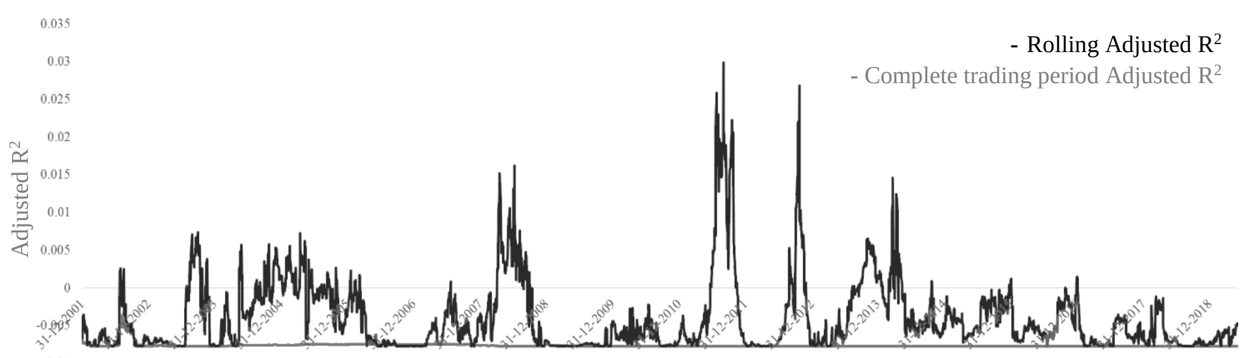
Appendix J – Cumulative Return for the Factor MKTRF with transaction costs



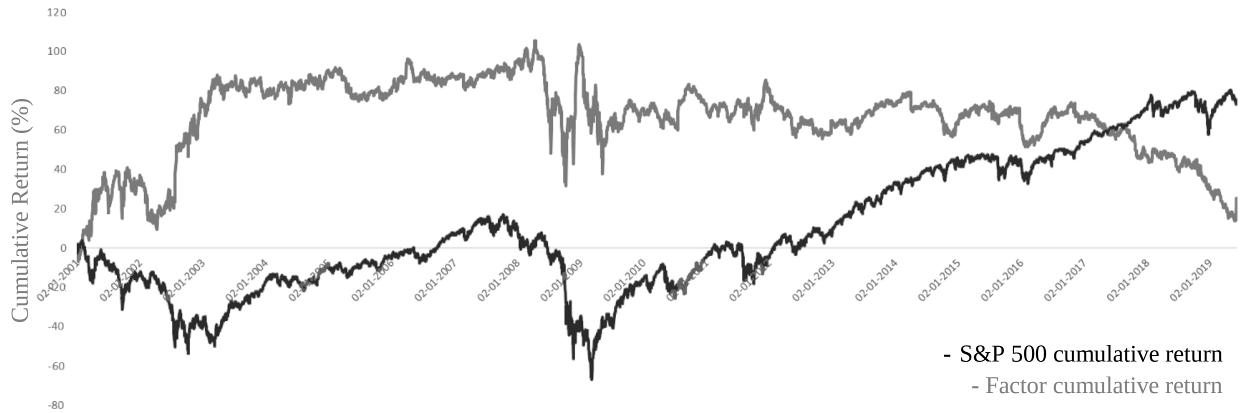
Appendix K – P-value of the Factor MKTRF with transaction costs



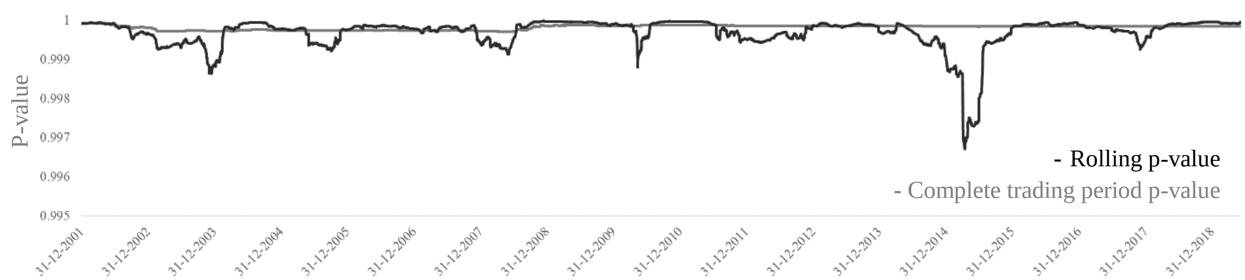
Appendix L – Adjusted R Squared of the Factor MKTRF with transaction costs



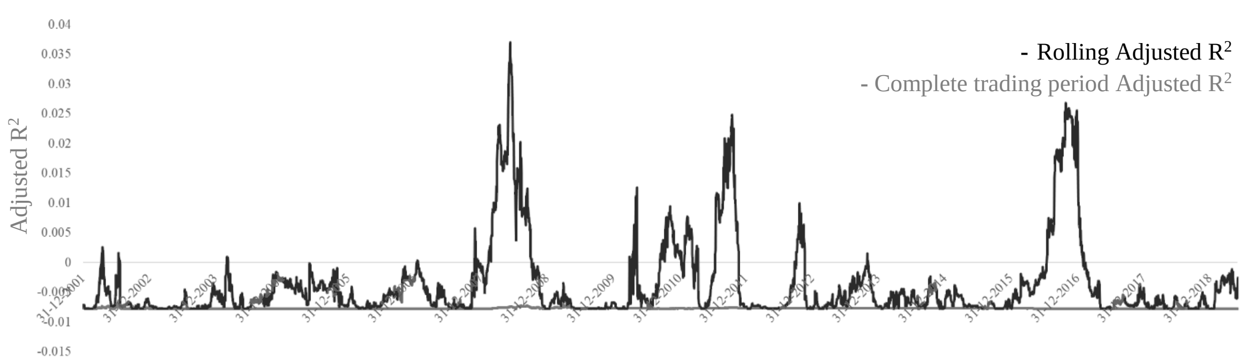
Appendix M – Cumulative Return for the Factor MOM with transaction costs



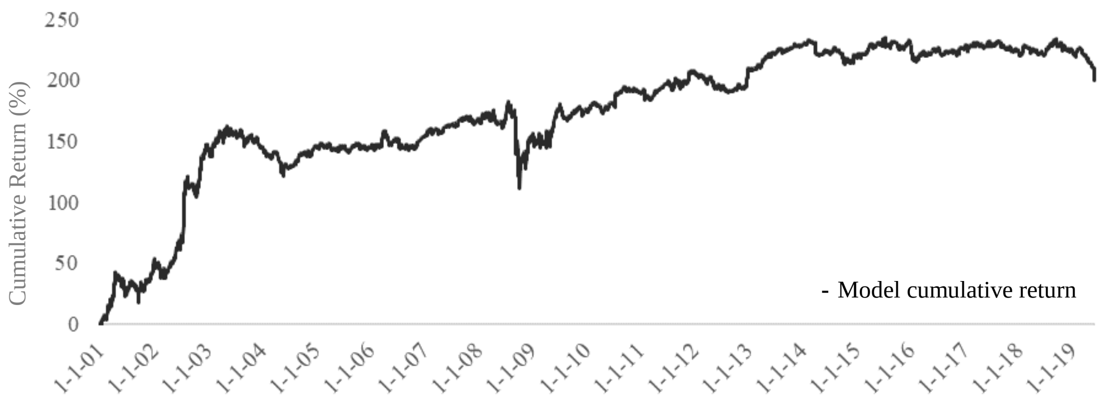
Appendix N – P-value of the Factor MOM with transaction costs



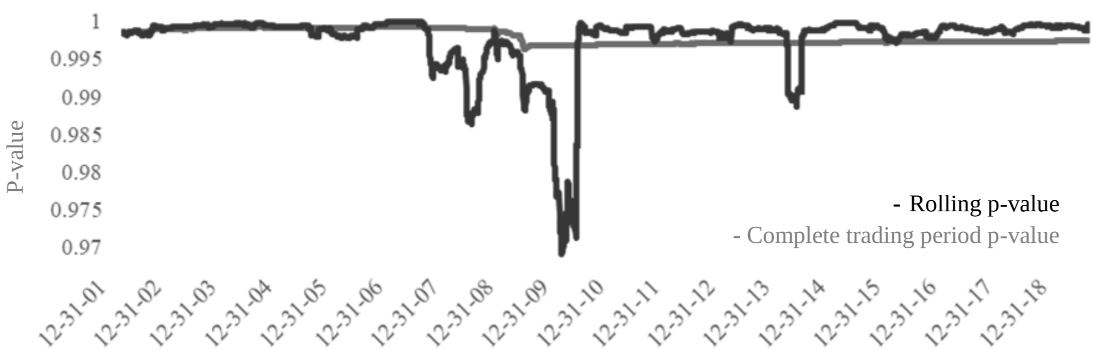
Appendix O – Adjusted R Squared of the Factor MOM with transaction costs



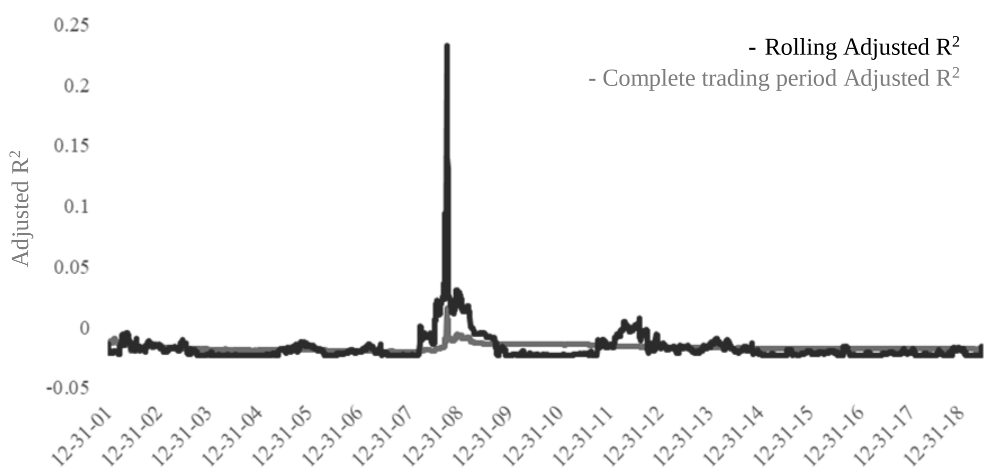
Appendix P – Cumulative Return for the Full Model without transaction costs



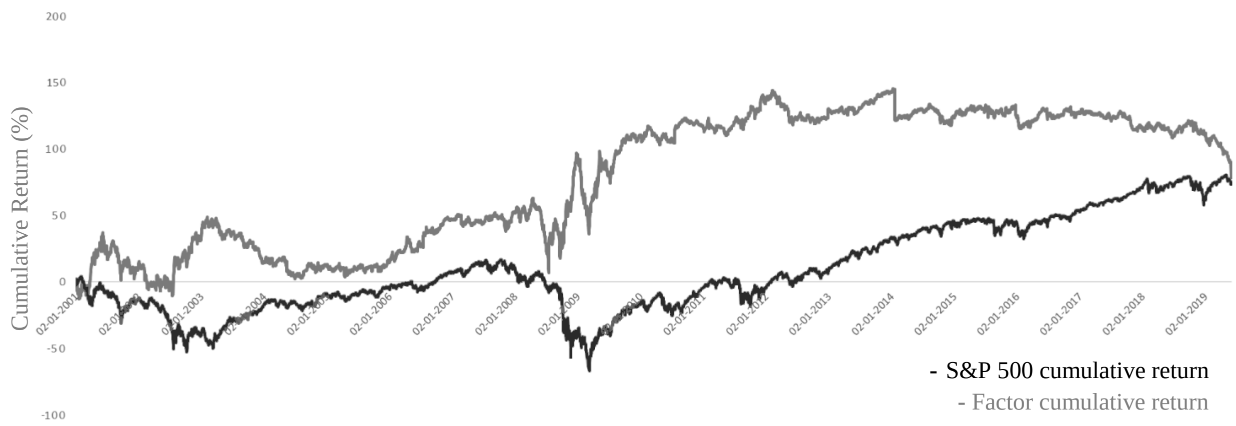
Appendix Q – P-value of the Full Model without transaction costs



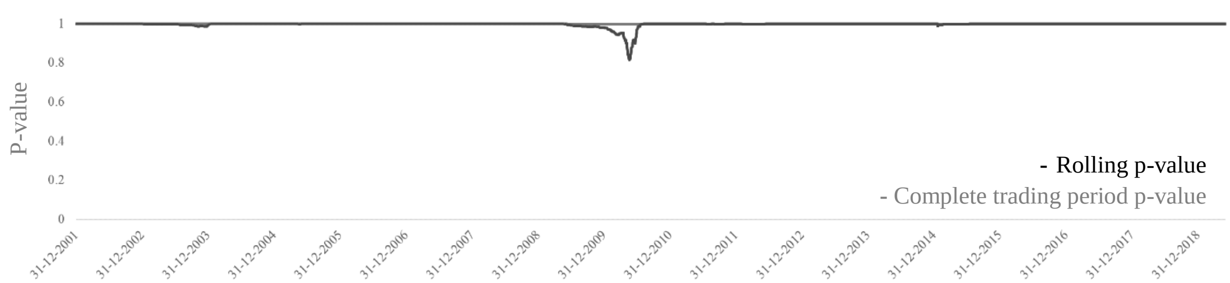
Appendix R – Adjusted R Squared of the Full Model without transaction costs



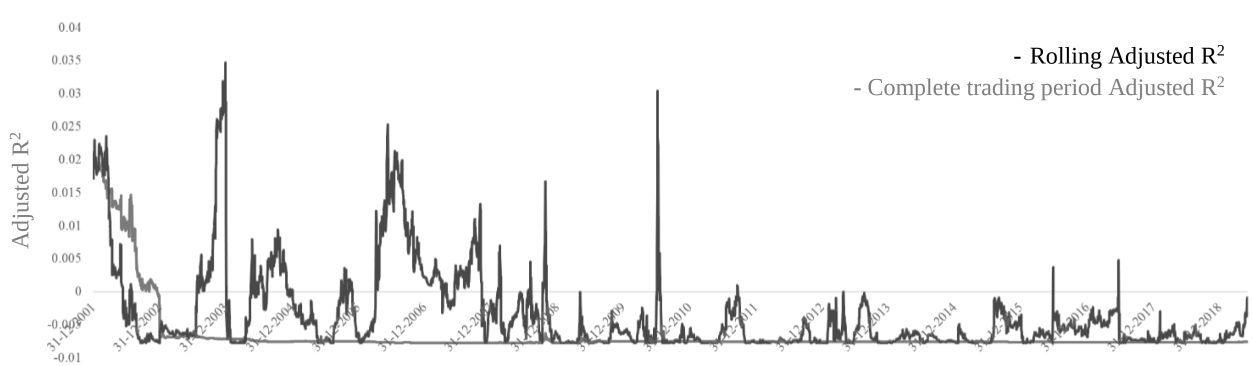
Appendix S – Cumulative Return for the Factor Sharpe without transaction costs



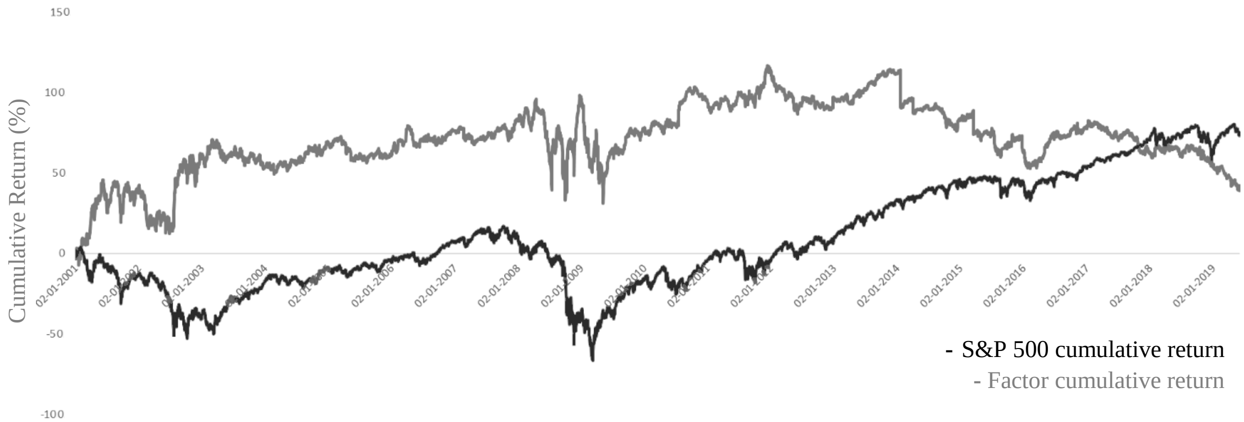
Appendix T – P-value of the Factor Sharpe without transaction costs



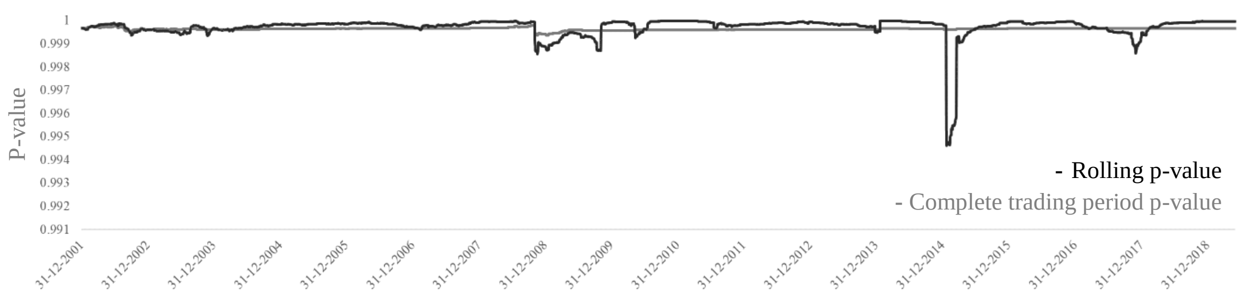
Appendix U – Adjusted R Squared of the Factor Sharpe without transaction costs



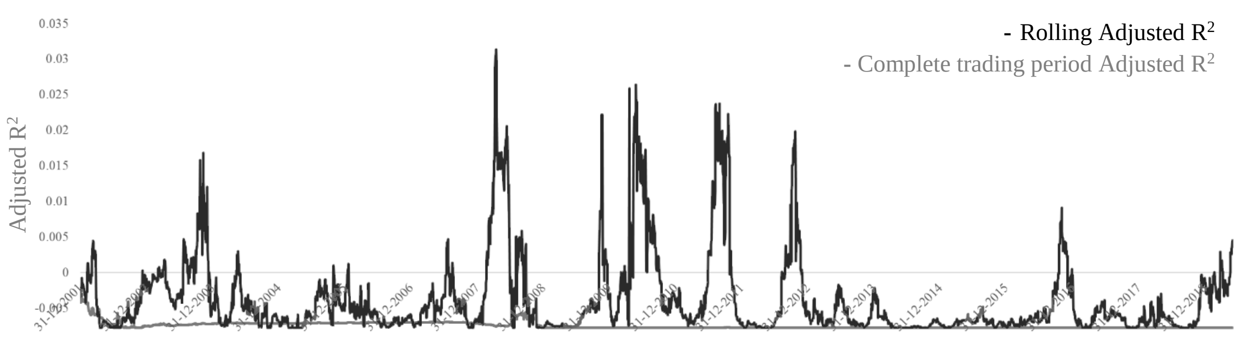
Appendix V – Cumulative Return for the Factor SMB without transaction costs



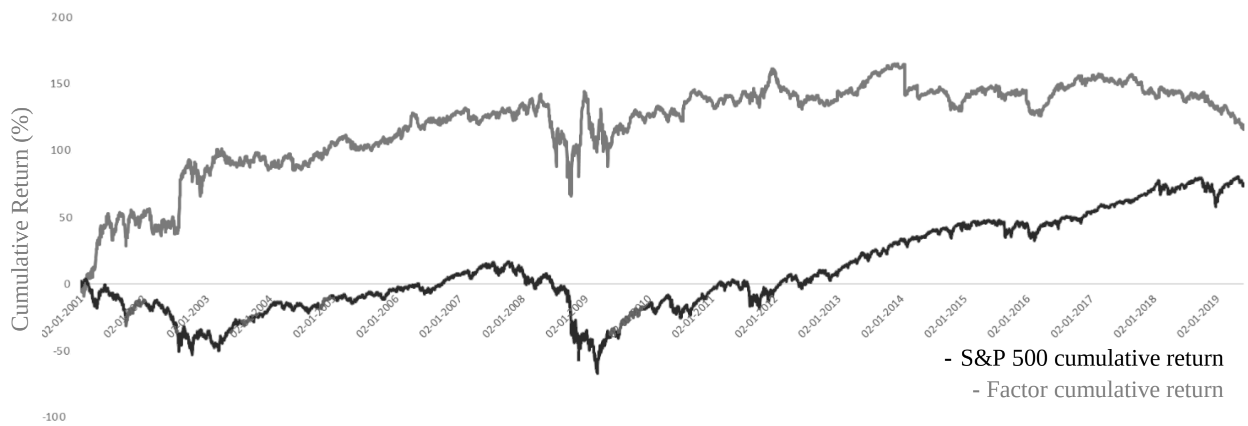
Appendix W – P-value of the Factor SMB without transaction costs



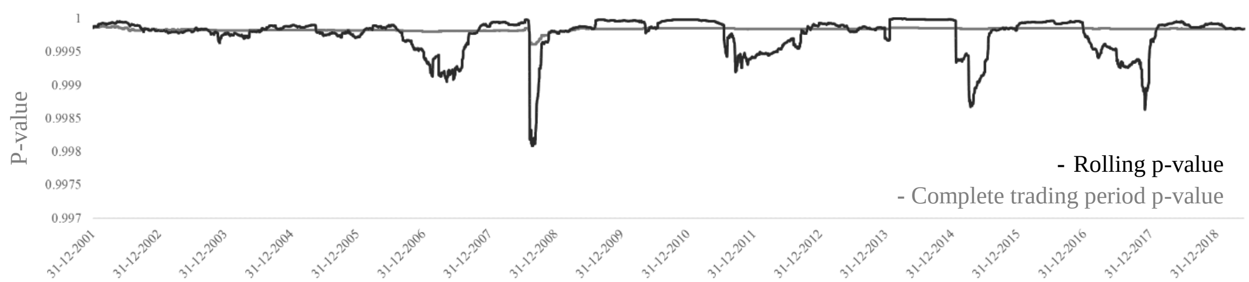
Appendix X – Adjusted R Squared of the Factor SMB without transaction costs



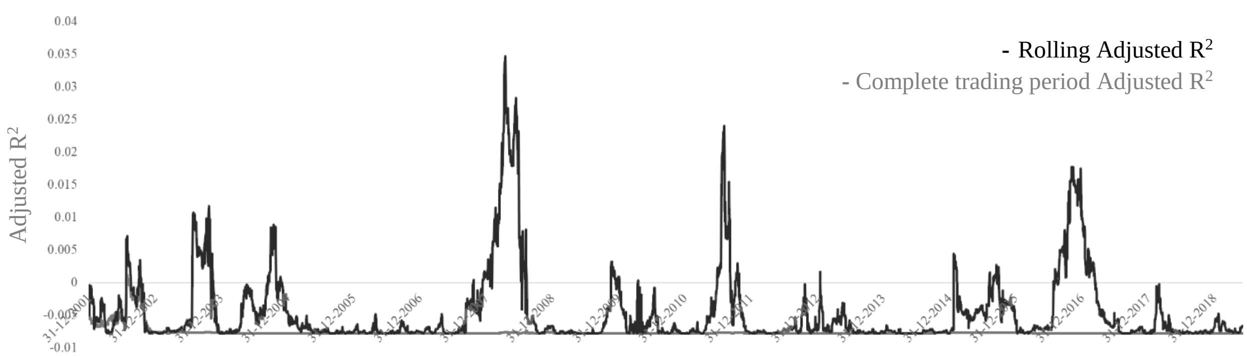
Appendix Y – Cumulative Return for the Factor HML without transaction costs



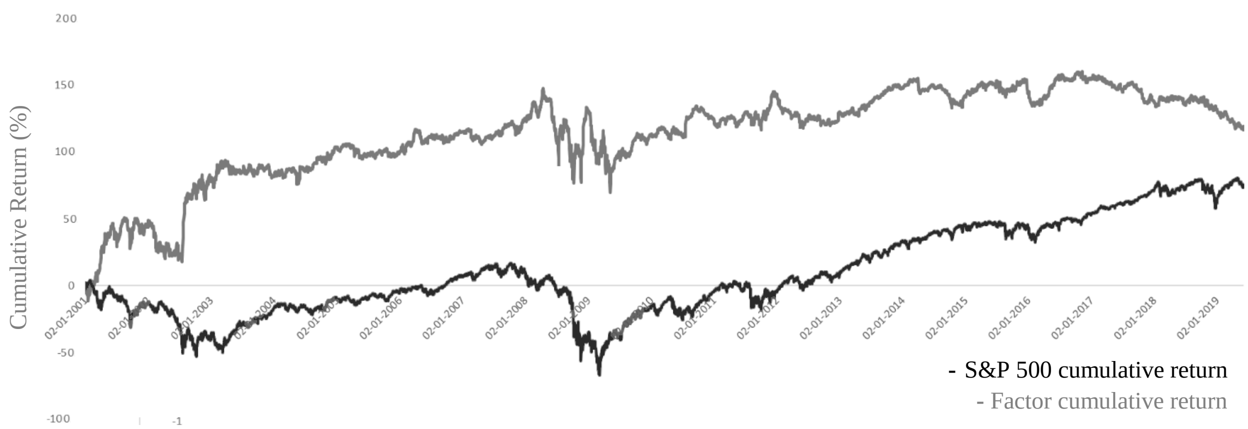
Appendix Z – P-value of the Factor HML without transaction costs



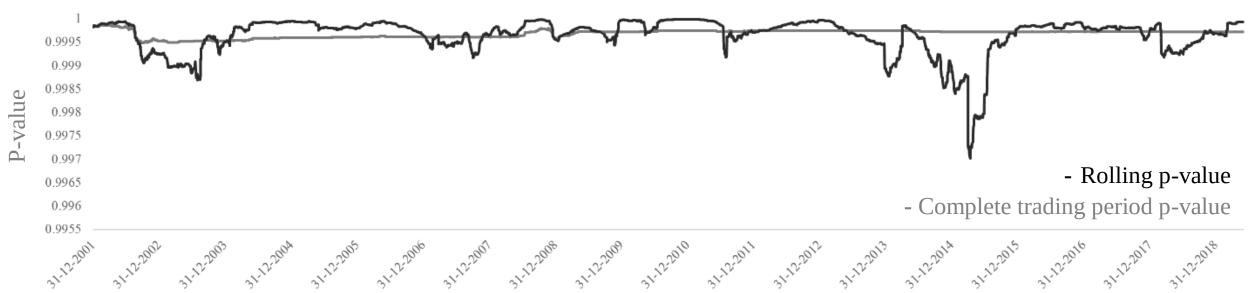
Appendix AA – Adjusted R Squared of the Factor HML without transaction costs



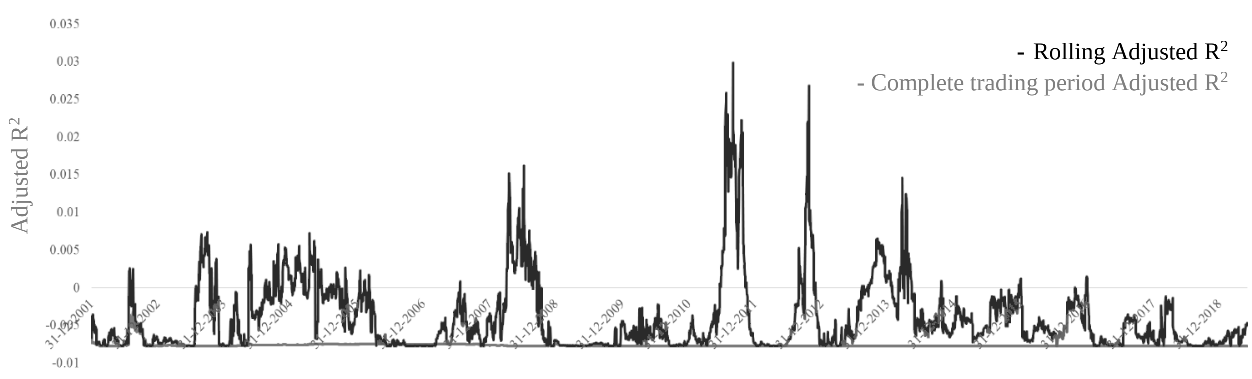
Appendix AB – Cumulative Return for the Factor MKTRF without transaction costs



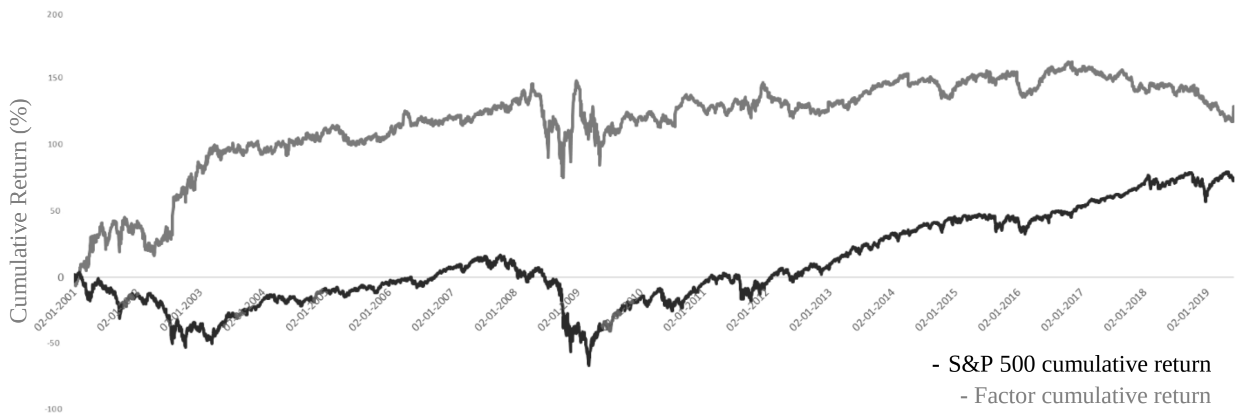
Appendix AC – P-value of the Factor MKTRF without transaction costs



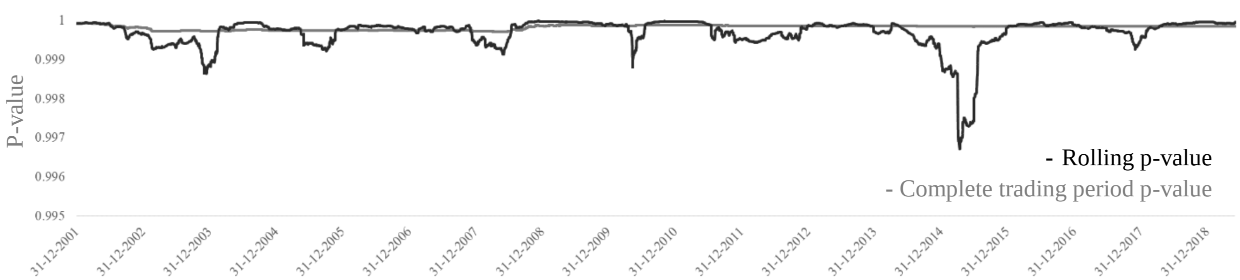
Appendix AD – Adjusted R Squared of the Factor MKTRF without transaction costs



Appendix AE – Cumulative Return for the Factor MOM without transaction costs



Appendix AF – P-value of the Factor MOM without transaction costs



Appendix AG – Adjusted R Squared of the Factor MOM without transaction costs