Predictability of Mutual Fund Performance

Evidence from the Swedish Market

Stefano Pagliani❖ Philipp Rusch♠

Master Thesis
Department of Finance
Stockholm School of Economics

May 2017

Abstract

This paper studies the out-of-sample performance of three established mutual fund performance predictors in Sweden. Our study draws inspiration from the work of Jones and Mo (2016). We analyse the momentum, back testing and selectivity predictors, based on previously introduced methodologies, as well as our own revised approaches. Our results only partly confirm those of the original studies, as the predictors exhibit diverse performance. While the momentum predictor underperforms three Swedish equity benchmark indexes, the back testing predictor exhibits unclear results. The selectivity predictor outperforms the other two methodologies and all our benchmarks. Our results are gross of all fees and expenses. Estimated alphas from Carhart (1997) four-factor regressions are positive and significant for the modified back testing predictor and for both approaches of the selectivity predictor. The results are potentially important for investors in Swedish mutual funds.

Keywords: Performance Prediction, Investment Decisions; Mutual Fund Performance
JEL classification: G11, G14, G23
Supervisor: Michael Halling[1], Associate Professor Department of Finance

[1] We are thankful for the continuous valuable support and guidance during the writing of our thesis
❖ 40937@student.hhs.se
♠ 40951@student.hhs.se

Table of Contents

List of Figures

List of Tables

1. Introduction

The mutual fund industry has grown to vast size over the past decades. This naturally draws attention from the academic community, which has addressed many important characteristics of mutual funds to this day. One field of interest deals with the question of whether investors can profit from identifying particularly skilled mutual fund managers. A large set of academic papers has implemented various statistical tools, based on specific mutual fund characteristics, to investigate the matter, in part with respectable success. Even though literature seems to agree that mutual funds, on average, do not outperform passive benchmark indexes (e.g. Sharpe (1966), Jensen (1968) or Treynor and Mazuy (1966)), several papers have outlined short-term persistence patterns in mutual fund returns. Grinblatt and Titman (1992) find positive persistence in mutual fund returns, which, they argue, cannot be explained by inefficiencies of their model. Other studies finding performance persistence include Hendricks et al. (1993) and Carhart (1997), who also introduce methodologies to capitalise on this persistence. Many papers have tried to predict future fund performance, partly using proprietary statistical tools, so called mutual fund performance predictors. The predictors make use of various fund characteristics, such as fund managers' stock selectivity, liquidity of fund holdings or fund size. Jones & Mo (2016) provide a comprehensive view of the most important of these studies and test their out-of-sample performance. Their study focuses on the sample period and investigates whether fund performance predictors lose efficacy after they have been introduced by academia.

Our research draws inspiration from their paper, as we introduce another out-of-sample performance test of some of the predictors treated by Jones & Mo (2016). Naturally, most literature about mutual funds focuses on the US mutual fund industry, as this represents by far the biggest and, hence, most important market. There are, however, other markets with interesting mutual fund market characteristics, such as Sweden, which is the geographic focus of this study. We focus on three performance predictors: first, we test the performance of a momentum predictor, based on a paper by Hendricks et al. (1993) that takes into account a fund's recent performance. Second, we analyse the back testing predictor, which in fact represents a general improvement of momentum-based predictors and which was introduced by Mamaysky et al. (2007). Finally, we implement a predictor by Amihud and Goyenko (2013) that is based on fund selectivity measured by a fund's resemblance to benchmark indexes. We ultimately assess the performance of each predictor by running Carhart (1997) four-factor regressions.

The paper at hand differs from Jones and Mo (2016) in several ways. First, we focus on the performance of three predictors in the Swedish mutual fund industry instead of the US market. Second, not only do we investigate the efficiency of the original papers' methodologies, but also introduce our own modifications, which exhibit diverging results. Third, we examine the predictors' relative performance relative to each other and try to identify the most profitable predictor from an investor's point of view, even though we do not account for fees or expenses.

The findings of this study are interesting for the following reasons: first, the Swedish mutual fund industry, totalling SEK 2.427 trillion in 2016, is unique due to its link to the Swedish pension system, which implies that employees must invest part of their pension payments in a set of mutual funds. Employees have the choice between a total of 700 mutual funds and may invest in up to five funds. Therefore, investigating the performance of mutual fund performance predictors in the Swedish market could provide valuable insights and decision guidance for a great number of pension policy holders and other retail investors. Second, our findings could help assess whether it is worthwhile for investors to invest in more expensive actively managed funds compared to cheaper passive and index tracking funds or exchange-traded funds, which have increasingly challenged the actively managed fund sector in recent years. Third, if returns of trading strategies based on the presented predictors cannot be explained by the common risk factors market risk premium, size, value and momentum, there might be no well-functioning model for the mutual fund industry, which could prove to be an interesting field for future research.

The remainder of this paper is structured as follows: *Section 2* gives an overview of the existing literature about mutual funds and, more specifically, about mutual fund performance predictors. *Section 3* describes the data used and *Section 4* illustrates the methodologies of the predictors as well as statistical properties of our data. *Section 5* presents the results of our performance analysis. *Section 6* provides a critical review of our results, including robustness issues and suggestions for future research. Finally, *Section 7* concludes our paper.

2. Related Literature

The following section provides an introduction to the existing literature about mutual funds' performance. The second part focuses on literature about performance prediction, following the structure presented by Jones and Mo (2016).

2.1. Evidence on Mutual Fund Performance

During the past decades, an extensive body of literature about mutual fund performance has been established. Since the 1960s, researchers have attempted to answer the question of whether mutual funds can systematically generate positive alpha, i.e. outperform a passive market benchmark. Over the years, studies have been repeated with extended datasets and altered methodologies. Literature has brought forth contradictory evidence, but seems to agree at large that fund managers do not systematically outperform benchmark indexes after fees and expenses. During the 1990s, a substantial subarea of mutual fund performance research emerged: mutual fund performance predictors. Since then, an increasing number of publications has tried to use mainly publicly available fund characteristics and information to sort and pick mutual funds ex ante, in order to achieve positive alpha, largely with considerable in-sample success. This section aims at building the context for our study by providing an overview of the existing literature on mutual fund performance and performance predictors.

Research during the 1960s built the groundwork for literature about mutual fund performance. Treynor and Mazuy (1966) test whether mutual fund managers can successfully predict future market movements and find no evidence for that. Sharpe (1966) bases his study on new findings in the fields of portfolio analysis and asset pricing theory and argues that good portfolio managers focus on evaluating risk and diversifying their portfolios instead of searching for mispriced securities. He provides evidence that mutual fund performance from an investor's point of view, i.e. after fees and expenses, falls short of the performance of benchmark indexes. Further evidence that mutual funds do not, on average, outperform a market portfolio stems from Jensen (1968). In his paper, he uses a risk-adjusted measure of portfolio performance – Jensen's α – to investigate the performance of mutual funds compared to the Standard & Poor's 500 index, finding little evidence for significant outperformance of funds, both on average and on an individual fund level. It is worth noting that Jensen investigates mutual fund performance, both gross and net of fees, reaching the same conclusion in both cases. Other studies report similar findings: Malkiel (1995) finds that mutual funds underperform the market, both net and gross of fees. However, he also outlines, at least during the 1970s, performance persistence patterns at a fund level that he cannot detect during the 1980s. Grinblatt and Titman (1989a) use Jensen's measure and find a number of funds, namely aggressive growth funds and mutual funds with lower net asset values (NAV), that outperform market benchmarks. These findings are, however, limited to gross returns as the identified funds exhibit the highest fees in their sample, thereby eating up profits for investors. Barras et al.

(2010) even find that a substantial portion of mutual funds overcharge investors relative to the skills of their managers, i.e. alphas net of fees and expenses are significantly negative.

Nonetheless, several studies report findings of performance persistence, i.e. even though it seems that fund managers cannot systematically outperform market benchmarks, they might be able to do so for a certain period of time. Grinblatt and Titman (1992) find that differences in fund performance are persistent over a five year time horizon. Hendricks et al. (1993) provide evidence of short-term performance persistence – they find the strongest evidence for a period of twelve months. Carhart (1997) also finds short-term performance persistence, even though somewhat weaker than other studies, and explains most of the resulting spread between good and poor performers by common risk factors for stocks as well as fund expenses and transaction costs. Further evidence stems from Kosowski et al. (2006), who find that a very small number of mutual fund managers produce superior alphas that persist during a three year time window.

2.2. Mutual Fund Performance Prediction

Regardless of the question of whether fund managers can systematically outperform the market, a variety of studies have found that certain variables, i.e. some mutual fund characteristics, can be used to predict future mutual fund performance. Most of the studies were published with data derived from the newly available CRSP Survivorship Bias Free US Mutual Fund Database. Jones and Mo (2016) identify a total of twenty different mutual fund performance predictors from seventeen papers over time. We present the most important of these studies hereafter.

Hendricks et al. (1993) introduce the first of these papers, finding short-term persistence in mutual fund performance and using it to predict future "winners and losers" amongst funds. Specifically, they rank mutual funds by lagged one-year excess returns and group them into octiles, to either go long the octile of winner funds or to go long the octile of winner funds and short the octile of loser funds. However, only the portfolio including the biggest underperforming mutual funds exhibits statistical significance.[2] Even though the study predates the introduction of the above mentioned CRSP database and was challenged as being biased by survivorship (Malkiel, 1995), Carhart (1997) confirms the results of Hendricks et al. (1993) using the CRSP database. Furthermore, Carhart (1997) finds that turnover and expense ratios substantially lower fund performance. Funds charging loads significantly underperform funds that do not charge such fees. Other studies that employ approaches similar to the ones above

[2] Throughout our study, we define statistical significance as significance at a confidence level of 5%.

include Grinblatt and Titman (1992) and Goetzmann and Ibbotson (1994), who also present evidence for short-term performance persistence.

Another fund characteristic that arguably predicts mutual fund performance is fund size. Chen et al. (2004) find that both fund size and fund family size are related to fund performance. They demonstrate that fund size erodes fund performance and argue that this adverse effect arises mainly from costs linked to liquidity constraints. However, controlling for fund size, fund family size is positively linked to fund performance.

A large field of alpha predictors circles around the stock picking ability of mutual fund managers, thereby directly addressing fund managers' capabilities. Cohen et al. (2005) compare a fund manager's portfolio to the portfolios of other well-established fund managers with striking performance records. High portfolio resemblance is shown to go hand in hand with greater managerial skill and better fund performance. Kacperczyk et al. (2005) argue that mutual funds with more concentrated portfolios perform better than their counterparts, supporting the view that fund managers make use of their individual industry knowledge. Regressing changes in a fund manager's portfolio holdings on changes in public information, Kacperczyk and Seru (2007) find that more capable fund managers react less to public information. Additionally, they establish a negative link between a fund managers' reactiveness to public information and other fund performance predictors as well as fund flows. Christoffersen and Sarkissian (2009) follow a somewhat different approach and investigate fund managers' stock picking ability by looking at a fund's city size as indicator. They conclude that funds situated in large financial hubs perform substantially better than funds in smaller cities, mainly due to more available information and information spill-over effects. They also remark that fund managers in larger cities partly show patterns of performance improvements over time that cannot be observed in funds from small cities.

A similar group of mutual fund performance predictors identified by Jones and Mo (2016) measures a fund's deviation from benchmark indexes to draw inferences on its future performance. Cremers and Petajisto (2009) measure the divergence of mutual fund holdings from the composition of a benchmark index, mainly using the newly introduced Active Share measure. Funds with high scores in this measure, i.e. active stock picking funds, are shown to outperform mutual funds that exhibit lower scores of the same measure. A second predictor from this group stems from Amihud and Goyenko (2013), who use the R^2 from a regression of fund returns on the returns of a benchmark index. A high R^2 denotes high similarity to an index and implies worse future performance compared to funds with a smaller R^2, i.e. with greater divergence. The latter predictor clearly has the advantage that it does not require detailed fund holdings data for its computation, as such data may be difficult to obtain in a reliable and timely fashion, if at all. Further studies in this area include Brands et al. (2005), who find that deviation from benchmark predicts mutual fund performance for Australian mutual funds, and Daniel et al. (1997), who find that the average actively managed mutual fund outperforms mechanic strategies, even though the difference is usually equal to the average management fee.

Further predictors compare the realised performance of mutual funds with the performance of the assets they hold. Kacperczyk et al. (2008) observe the difference between fund returns and the actual returns of their holdings by building replicating portfolios. Their reasoning behind this approach is that not all actions taken by mutual fund managers are revealed, as holdings are disclosed with set frequencies, which can be monthly, quarterly or even annual. They find that these undisclosed actions create a "return gap" that predicts mutual fund returns. Huang et al. (2011) investigate risk shifting of mutual funds, i.e. loading on large amounts of risk in times of distress, and find that it has a negative effect on fund performance. Funds that show high levels of risk shifting tend to underperform funds with constant risk-levels.

An important factor for mutual funds is the liquidity of their holdings. Da et al. (2011) make use of the PIN measure developed by Easley et al. (1996), which indicates the amount of information that affects a certain stock. They assess that superior fund managers benefit from trading high PIN measure stocks, i.e. stocks that are highly affected by information. Another predictor based on liquidity comes from Cao et al. (2013). They find that some mutual fund managers show the ability to time market liquidity, thereby performing better on average than their counterparts. More precisely, Cao et al. (2013) show that some fund managers increase their exposure to the market prior to periods of increased liquidity.

Finally, Jones and Mo (2016) identify two performance predictors that try to improve other methodologies. Mamaysky et al. (2007) introduce a back testing method that checks a fund's predictive power before it is used in the prediction and find that prior testing increases a predictor's efficiency. This procedure also seems to reduce the number of funds with incorrectly estimated parameters in the top and bottom portfolio, which are the decisive portfolios when it comes to constructing trading strategies based on the prediction. Kacperczyk et al. (2014) combine a market timing measure, used in periods of economic recessions, with a stock picking measure, which is used in economic boom periods. Based on this combination of measures, they argue that skilled mutual fund managers do exist and that these managers significantly outperform the market.

3. Data

For our analysis, we make use of four main data sources. First, we obtain a register of Swedish mutual funds from Finansinspektionen (FI), Sweden's financial authority. Second, we acquire daily, monthly and quarterly NAVs for the mutual funds from Bloomberg. We retrieve data for the risk-free rate and the US Dollar – Swedish Krona exchange rate from the Swedish Riksbank and use Thomson Reuters Datastream to obtain data for the benchmark indexes.

Data from FI comprises quarterly holdings as disclosed by individual funds to the authorities.[3] The oldest data dates to the first quarter of 2003 and goes until the last quarter of 2016. The initial data includes a total of 1,043 different funds. We clean the data according to different requirements, aiming at a database consisting solely of actively managed funds that invest predominantly in Swedish equity. Hence, we analyse the holdings and exclude a fund quarter if the share of Swedish equity held by the fund within the quarter falls short of our threshold of at least 70% of total assets. Furthermore, we examine the funds by scanning the fund names for components signalling that the fund does not fulfil our requirements of active management and geographical focus on Swedish equities. Accordingly, we exclude funds from our sample if their names include certain keywords.[4] In order to reduce the risk of distortion through outliers, we exclude funds that have ten or fewer holdings or report for less than four quarters. Additionally, we double check each fund in our sample by researching its ISIN and its description on either Morningstar or Bloomberg. In the end, our sample consists of 105 funds that we identify as actively managed Swedish equity funds, fulfilling all our requirements. Compared to Jones and Mo (2016), our study is based on a much smaller dataset. Since they focus on the US market, which entails a much higher number of equity funds and a more extensive data record, this is to be expected. Unfortunately, holdings data provided by FI is not complete. In particular, the data is somehow fragmented: no holdings are reported for the first, third and fourth quarter of 2004 and the first three quarters of 2005. We try to account for this by checking each fund's description, as explained above. We analyse each instance and decide on a case-by-case basis up to what date we include the fund in our sample. Our goal is to

[3] The Swedish financial authorities unfortunately no longer report fund holdings. Our data package is dated January 2017 and contains all holdings until the end of 2016.

[4] We exclude funds such as hedge funds, exchange traded funds, index funds or passive funds i.a. by searching for the following keywords: "etf; exchange-traded; exchange traded; hedge; index; leverage; levered; OMX; OMSX; passive; passiv". We also search for keywords that indicate that a fund is mainly invested internationally, which would also disqualify it for our study. Such keywords are: "America; Asia; Dow Jones; Europe; international; NASDAQ, s&p; s & p; s & p 500; s&p500; s&p 500; US; USA; world". Funds identified through these keywords are scrutinised based on their holdings reported to Finansinspektionen.

maintain continuity, so we assume the fund satisfies our requirements if data is missing for 1-4 quarters.

After filtering the funds by our described requirements, our sample includes 10,639 months in total, with an average of 101 months and a median of 105 months per fund. 36 funds persist for the whole sample period, while two funds only exist for twelve months. In two instances, we do not have data for the return of the fund. In these cases, we back out the monthly return from the quarterly return.[5] We retrieve daily NAVs for each of the filtered funds from Bloomberg. We then compute daily, monthly and quarterly returns from the daily NAVs. We calculate returns on a NAV to NAV basis, with gross dividends re-invested on ex-dividend date. *Table 1* reports summary statistics of our filtered data. Fund monthly excess returns range from -23.47% to 33.52%, with an average return of 1.00%. The sample is quite variable, with a standard deviation of 5.18%.

Final Sample Summary Statistics

	12 ≤ Months ≤ 60	60 < Months ≤ 120	120 < Months ≤168	Fund-month Observations
No. of funds	38	21	46	10,639

	Minimum	Maximum	Mean	SD
Funds' monthly excess returns	-23.47%	33.52%	1.00%	5.18%

Table 1 - Summary statistics of the mutual funds in our final sample
No. of funds indicates the number of funds that are present and report returns for the corresponding number of months. Fund-month observations states the total number of monthly returns present in our sample. Funds' monthly excess return summarises the distribution of the fund returns in excess of the risk-free rate, including its minimum, maximum, mean and standard deviation.

We obtain data for the benchmark indexes from Thomson Reuters Datastream. We have identified the OMX Stockholm 30 (OMXS30) and the OMX Stockholm Mid Cap (OMXSMCPI) as appropriate benchmarks for our studies. The OMXS30 contains Sweden's 30 largest public companies, while the OMXSMCPI is comprised of 100 Swedish mid cap companies and represents a high growth index throughout our sample period. From the AQR database that we use for the Carhart four-factor regression, we derive a third benchmark, which is the market factor in those regressions and which we refer to as AQR Market Index (AQR Capital Management, LLC, 2017). Consistently with the data from FI, we obtain data for the period from 2003 through 2016 (168 months), computing monthly, quarterly and yearly returns from the daily data.

[5] November 2013 for fund 51308 and June 2015 for fund 51437.

Data on risk free rates is obtained from the Swedish Riksbank. Again, we retrieve daily, monthly and yearly frequencies. We adjust the rate from the one month treasury bill to the desired frequencies using simple compounding. Furthermore, we obtain data on the USDSEK cross rate from Riksbank. The data has daily frequency and spans our whole sample, from 01/01/2003 to 31/12/2016. We use the FX rate to convert returns stated in USD into SEK returns, as described below.

We get data for the Carhart four-factor model from the AQR database (AQR Capital Management, LLC 2017). The model is based on four portfolios, which represent four common risk factors: market portfolio excess return, high-minus-low (HML) book-to-market ratio stocks, small-minus-big (SMB) size stocks and, finally, winner-minus-loser (UMD) stocks. AQR determines the four factors following an approach by Asness & Frazzini (2013). We use said database as it provides us with daily data for Sweden, whereas Kenneth French's website only provides monthly data for Sweden. Data from the AQR database is expressed in USD and, consistently, in excess of the US risk-free rate for the measures that require so. Thus, we follow steps to make this data appropriate for our study. First, we add back the US risk-free rate, included in the AQR database. We then adjust for the change in the USDSEK cross rate to derive the implied return of the Swedish market[6]. Finally, we subtract the Swedish risk-free rate to reach the market excess return. The HML, SMB and UMD factors are merely converted in SEK returns.

Due to the fact that fund returns from Bloomberg, the Carhart four-factors and the risk-free rate from Riksbank all have some differences regarding reporting days, we base our data on the Swedish trading calendar. As such, we match data from the other datasets to the reporting days from Bloomberg. We interpolate risk-free rates in the instances where we do not have data. This is not necessary for the Carhart four-factors, as this dataset includes all required dates of the Swedish trading calendar. Overall, we are confident about the quality of the data we use throughout the study. The databases we base our research upon are all reputable. However, it should be noted that the data provided by FI might contain some erroneous reporting, as we received communication from the institution that data may not be reliable. This possible flaw is limited to the quality of data about fund holdings. We approach the problem by manually checking each fund's description on Bloomberg or Morningstar. As a precaution, we refrain from replicating predictors that are based on holdings. Related studies are almost exclusively based on the US market and, hence, make use of the CRSP Survivorship Bias Free US Mutual

[6] We adjust using the following formula: $SEK\ return = (1 + USD\ return) * (1 + USDSEK\ return) - 1$.

Fund Database. Since a comparable database for the Swedish market does not exist, we try to mitigate the survivorship bias manually. We stop recording data of a fund in our sample only in case the fund ceases to exist, stops fulfilling our holdings criteria or merges with another fund. All in all, we believe that these measures work reasonably effectively.

4. Methodology

In this paper, we analyse different predictors to investigate whether they can successfully predict mutual fund performance in Sweden. We focus on three predictors out of the twenty that Jones and Mo (2016) analyse: a *momentum predictor*, based on the lagged one-year return predictor first introduced by Hendricks et al. (1993), a *back testing predictor*, developed by Mamaysky et al. (2007) and, finally, a *selectivity predictor*, analysed by Amihud and Goyenko (2013). Including more predictors would go beyond the scope of this paper. We follow Jones and Mo (2016) in their approach of introducing and analysing a set of predictors. Even though their paper serves as inspiration to this study, our methodology is based on the original predictor studies. Ultimately, they compare their out-of-sample results to the original in-sample performances. Our comparison is somewhat more direct, as we try to identify and point out the best performing of the predictors we implement.

The selection of those specific predictors is, mainly, due to unavailability of data, needed to construct most of the other predictors, such as *Expenses, Manager Tenure, Style*. This data is available for funds in the US but no such database exists for funds in Sweden, at least not to our knowledge. Additionally, there are instances when data is available, but not reliable. At the time of writing, *Holdings* data, retrieved from the FI, is described to us as uncertain, precluding us from analysing all holdings-based predictors. Other predictors are excluded as they would make little sense in the Swedish context; i.e. location-based predictors. However, strong motivation to focus on the three predictors we choose comes from the fact that all three are, in theory, replicable by retail investors, opening the door to trading strategies based on the results of our study. Contrary to Jones and Mo (2016), we do not have “in-sample” performance for predictors, as the original papers analysing such variables focus on US data. Furthermore, they investigate the effect of the publishing of the original studies on the effectiveness of the predictors.

4.1. Momentum Predictor

Hendricks et al. (1993) address short-run performance persistence and introduce a performance predictor that can be loosely described as a momentum predictor, as it looks at recent past performance to identify funds that should outperform relative to the others. Their study is successful in finding such a positive relation and in identifying funds with a positive or negative return streak ex ante. Thus, Hendricks et al. (1993) show that serial correlation of mutual funds' excess returns creates the opportunity for economically worthwhile fund selection strategies.

The study uses quarterly mutual fund returns, adjusted for fees and expenses but not for front- or back-end loads, over the period 1974 to 1988. This data is taken from the CRSP Survivorship Bias Free Mutual Fund Database. The risk-free rate is the return on one-month U.S. Treasury bills over the quarter. They identify 165 funds that satisfy their requirements. Three sets of benchmarks are used in their paper: single portfolio benchmarks commonly used in performance and equity pricing studies, an eight-portfolio benchmark, called P8, derived from Grinblatt and Titman (1989b), and an equally weighted index of mutual funds in the sample.

While the original study is quite extensive, especially regarding the proof of statistical autocorrelation in mutual fund returns, we focus on the performance potential of trading strategies that exploit the discovered short-term performance persistence. To do so, we construct the momentum predictor closely following the procedure of Hendricks et al. (1993). Thereby, we can assess if those exploitation strategies yield economically meaningful benefits in the setting of the Swedish mutual fund industry. However, we also run statistical tests on the final portfolios. More details about these tests can be found in *Section (4.4.)*. We follow two different approaches, one with quarterly data, as used in the original study, and one with monthly data, which represents our modified approach.

The portfolios are built as follows: for every period in our sample, which can be one quarter or one month, depending on the approach, we distribute funds in eight performance-ranked portfolios. The ranking is based on the funds' performance over an estimation window. This estimation window is variable, although Hendricks et al. (1993) show it to be optimal, in terms of alpha, at a one year length. We also investigate the effects of variations in the estimation window, which span from one quarter to two years and find the one year to yield the best results. Accordingly, our results are derived from a one year estimation window. The ranking and allocation process follows certain conditions that must be met in all cases.[7] *Appendices 1* and *2*

[7] In period t, order the N_t available funds based on their excess return over the estimation window. Let the rank of fund I's return be $rank(r_i)$. The fund is assigned to octile j such that the following formula is satisfied:

report summary statistics of the sorting variable, which is the lagged performance of funds, for the construction of the portfolios. We compute the average of the lagged performance of funds that constitute a portfolio for each portfolio in each period. *Appendix 1* refers to the original approach, i.e. with quarterly data and a holding period of one quarter, while *Appendix 2* refers to our modified approach, with monthly data and a holding period of one month. We can see that the sorting works as expected, with mean performance increasing in octiles. This observation holds for both the original and modified approach. The original approach reports mean performance ranging from 0.89% to 5.47% quarterly, while the range is 0.28% to 1.71% monthly for the modified approach. Minimum and maximum values are also increasing in octiles, although not monotonically. This suggests that the ranking procedure is not entirely accurate. The standard deviation of the average performance is constant at around 9% quarterly for the original and around 5% monthly for the modified approach. Converting the monthly standard deviation to quarterly frequency yields a value of 8.66%.

Next, we enter a long position in the 8$^{\text{th}}$ octile portfolio, which represents the portfolio composed of the best performing funds by construction. The holding period depends on the approach that we use and is always equal to one period, i.e. one quarter in the original approach and one month in the modified approach. After one period, the procedure is repeated, analysing the funds' performance in the past year, ranking them and building portfolios. Naturally, the number of funds per portfolio varies over time. We refer to the resulting portfolios as "octile portfolios". In our analysis, we focus specifically on the best performing portfolio, i.e. the 8$^{\text{th}}$ octile portfolio.

The performance of the portfolios of funds is evaluated during the holding period. If a fund is present during the evaluation period, but does not report during the following holding period, we assign it an arbitrary return of -20%. This way, we adjust for a potential hindsight bias, which would occur if we did not invest in a fund because we knew it would cease to exist in the following period. Once returns in all periods are obtained for all the funds, we run a linear regression of each portfolio's excess returns on the Carhart four factors:

$$\left(R_{j,t} - rf_t \right) = \alpha_j + \beta_{1,j}(RM_t - rf_t) + \beta_{2,j}HML_t + \beta_{3,j}SMB_t + \beta_{4,j}UMD_t + e_{j,t.} \qquad (1)$$

$R_{j,t}$ is the return of fund *j* in period *t*, from which we subtract rf_t, the risk-free rate for period *t*, to arrive to the fund's excess return for period *j*. The market factor is the AQR Market Index,

---

$$(j-1)\left[\frac{N_t}{8}\right] + \sum_{k=1}^{j-1} \delta_k \leq rank(r_i) < j\left[\frac{N_t}{8}\right] + \sum_{k=1}^{j} \delta_k \text{, } where\ \delta_k = \begin{cases} 1\ if\ k\ \leq N_t\ mod\ 8, and \\ 0\ otherwise. \end{cases}$$

while the other three factors are the usual HML, SMB and UMD, relative to the Swedish market. $e_{j,t}$ exhibits the residual term of the regression. We compute and report estimates for the portfolios' alphas and betas. As time-series regressions may be affected by autocorrelation and heteroscedasticity, we implement statistical tests to investigate these issues. Namely, we run tests developed by Durbin and Watson (1951) to check our data for autocorrelation and by Breusch and Pagan (1979) to test for heteroscedasticity. We refer to these tests as Durbin-Watson test and Breusch-Pagan test throughout the study. Furthermore, we make use of an estimator introduced by Newey and West (1987) (we refer to it as Newey-West estimator) and provide both adjusted and unadjusted t-statistics to account for potential statistical interferences. The estimator constructs a new variance-covariance matrix of the residuals that accounts for the issues of autocorrelation and homoscedasticity. More information about these tests can be found below in *Section (4.4)*. Finally, the cumulative performance of the 8th octile portfolios, both from the original and modified approach, is plotted against the three benchmarks.

4.2. Back Testing Predictor

Next, we turn to a predictor based on the approach by Mamaysky et al. (2007), who introduce the back testing methodology as an extension to earlier established momentum predictors, such as the one by Hendricks et al. (1993), with the goal of improving their predictive power. The paper argues that a traditional OLS model cannot consistently predict which portfolio managers will be able to produce above market returns. In particular, they reason that sorting on estimated alpha may populate the top and bottom decile portfolios with those funds which have the greatest estimation error. A poorly specified model will yield a wrong value for beta, which may be too low or too high. Consequently, the model will try to fit the return by raising (or lowering) alpha. To avoid this malfunction of momentum predictors, they introduce a back testing procedure as an extra step, which requires the model to show a certain degree of past predictive accuracy, prior to the actual estimation, before it is used to predict future performance. According to Mamaysky et al. (2007), the procedure yields an improvement of the risk-adjusted return for top mutual fund deciles. After the filtering step, funds are ranked by alphas.

Data used in the paper is taken from the CRSP Survivorship Bias Free Mutual Fund Database, complemented with the monthly CRSP value weighted return market index and the Fama-French and Carhart factors, provided by the WRDS web site. The sample spans from January 1970 to December 2002. Mutual fund returns are adjusted for fees and expenses but not for front- or back-end loads. We replicate the study, applying the back testing procedure to Swedish equity mutual funds. As our data sample is somewhat limited compared to the original one, both in terms of length and width, we adopt two different procedures. First, we follow Mamaysky et al. (2007) in using a five-year estimation window to calculate the estimated alpha. However, as that would leave us with only eight years of data to analyse (2008 through 2016), we introduce an alteration to the original approach, reducing the estimation window to one year. The alternative approach follows the same procedure using daily instead of monthly data.

The procedure is as follows: first, for each fund, we estimate the model with five years of monthly data or with one year of daily data up to time *t-2*, retaining this regression's alpha and beta estimates for *t-1*:

$$\left(R_{j,t} - rf_t \right) = \alpha_j + \beta_{1,j}(RM_t - rf_t) + \beta_{2,j}HML_t + \beta_{3,j}SMB_t + \beta_{4,j}UMD_t + e_{j,t.} \qquad (2)$$

The variables and factors in the regression are the same used in *Equation 1*. In *t-1,* we compare the sign of the estimated alpha to the sign of the fund's actual excess return over the market benchmark (the realised alpha). If the signs are the same, we add the fund to the active fund pool. Otherwise, the fund is left in the inactive fund pool for this month. As Mamaysky et al. (2007) point out, this procedure does not bias the estimated out-of-sample alphas towards high-risk funds, as the alphas are calculated from factor models that should eliminate such bias.

Next, we run the same model as (2) for all funds in the active pool using 60 months of data (or one year of daily data), until *t-1*, i.e. with a forward shift of one period compared to the previous step. The period by which we shift forward is always one month, for both approaches. Subsequently, a second filter is applied, using data from this second regression. Particularly, we check whether the new estimated alpha is between -0.02 and 0.02, and the estimated market beta is between 0 and 2, following the approach of the original study. We implement this filter to examine whether the obtained coefficients are reasonable and not outliers. In case those requirements are fulfilled, the fund's estimates for alpha and beta are retained and used to build portfolios of funds. Otherwise, the fund is moved to the inactive pool.

Funds in the active pool are sorted by alphas into deciles. We implement a reversed ranking, so that a lower rank implies a higher alpha and, supposedly, a future outperformance.[8] *Appendices 3* and *4* report summary statistics about the portfolios built following the original and modified approach, respectively. The tables refer to the alphas from the second regression

[8] This is due to the *rank* function in R, which starts from the first decile: in cases where the funds in the active pool are fewer than 10, the 10^{th} decile may not include any fund.

and present monthly and daily figures, respectively. We compute the average alphas of the funds that make up each portfolio, over time. Both sorting procedures work reasonably well. The 1st decile portfolio consistently reports the highest values, while the 10th decile portfolio presents the lowest ones. The original approach shows average monthly alphas that range from -0.00208 to 0.00505, while the range of daily alphas for the modified approach is from -0.00006 to 0.00073. Standard deviation of the alphas is fairly high, as it amounts to around 0.00160 for the original approach and 0.00046 for the modified approach. Comparing these values to the average alphas proves high volatility. We note that the modified approach leads to a much higher number of cases of no funds being in the active pool after the two steps. The number of instances amounts to 112 following the modified approach compared to only nine in the original approach. We assume an investment in the risk-free rate when no funds are in the active pool. As with the momentum predictor, we focus our analysis on the supposedly best portfolio, which is the 1st decile portfolio in this case. Returns for *t* are computed for each portfolio. Funds in the portfolios are held for one period (equal to one month), after which a new portfolio is constructed and invested in. We assume an equally-weighted investment in all funds. Following the same reasoning described in the momentum predictor, we assume a return of -20% for funds that do not report during a holding period.

The whole procedure is repeated every month. Once returns are obtained for all the funds in all periods, we follow the same steps that we follow when developing the momentum predictor: we run a linear regression of the portfolios' returns on the Carhart four-factors, report estimates for the portfolios' alpha and beta, run statistical tests and plot cumulative performance for the 1st decile portfolio. More information about the regression can be found above in *Section (4.1.)* and *Equation 1*.

4.3 Selectivity Predictor

Finally, we implement a selectivity predictor, derived from Amihud and Goyenko (2013). In their paper, they find that the R^2 obtained from a regression of a fund's performance on the Carhart four-factor model is a predictor of that fund's future performance. Specifically, they test the hypothesis that R^2 is negatively correlated to future alpha, the intercept from a Carhart four-factor regression, and to Info-Ratio, a measure derived from the Appraisal Ratio by Treynor and Black (1973). This hypothesis is consistent with the results of Cremers and Petajisto (2009) on selectivity. The reasoning behind this effect is that a higher R^2 implies a closer resemblance to the benchmark and thus lower selectivity. If we are to assume that selectivity improves mutual fund performance, R^2 should negatively predict a fund's future performance. The results from Amihud and Goyenko (2013) confirm this.

The data used in the original paper comes from the CRSP Survivorship Bias Free Mutual Fund Database and covers the period from 1989 to 2007. Returns are net of fees and expenses, but gross of front- or back-end loads. Daily data is retrieved from the International Center for Finance at Yale School of Management. This database provides daily data from 1989 to March 1998, after which daily CRSP data is used. We follow their approach when investigating the relation between R^2 and future alpha. Differently from their data, our returns are gross of all fees, as they are provided this way by Bloomberg. Additionally, we lack several of the control variables used in the original paper, such as *Expenses*, *Manager Tenure*, *Fund Age* or *Turnover*. Thus, we decide to run our regression without any control variables. Given the scarcity of our data set, we also introduce a modified approach. Instead of a one year holding period, we implement a one month holding period. This approach allows us to investigate the relationship between alpha or Info-Ratio and R^2 in greater detail.

Following the original study, we require a fund to have a minimum of 125 daily return data throughout the estimation year. For the following year, i.e. the estimation of a fund's performance measures alpha and Info-Ratio, we require at least 50 return data per fund. These requirements help to further reduce potential distortions arising from survivorship bias. For every year in the sample, we estimate the R^2 from an annual regression of daily fund excess returns on the Carhart four-factor returns:

$$\left(R_{j,t} - rf_t \right) = \alpha_j + \beta_{1,j}(RM_t - rf_t) + \beta_{2,j}HML_t + \beta_{3,j}SMB_t + \beta_{4,j}UMD_t + e_{j,t}. \qquad (3)$$

As in previous models, the regression tries to fit each funds' daily excess return on the Carhart four-factors (AQR Market Index, HML, SMB and UMD). See *Equation 1* for a more detailed description of each factor.

Next, we follow Amihud and Goyenko (2013) in applying a logistic transformation to R^2, which will make the resulting distribution more symmetric:

$$TR^2 = \log\left(\sqrt{\frac{R^2}{1-\sqrt{R^2}}}\right). \qquad (4)$$

We investigate the relationship between the transformation of R^2, TR^2, on alpha and on Info-Ratio. The Info-Ratio is a performance measure that considers potential survivorship bias by scaling alpha with the volatility of the abnormal returns. The Info-Ratio is computed as

$$Info - Ratio_j = \frac{alpha_j}{RMSE_j}, \tag{5}$$

with alpha being the alpha of the previously run Carhart four-factor regression and $RMSE_j$ being the squared root of the mean squared errors, or residuals, $e_{j,t}$ from the regression (3). We run a regression of funds' alphas or Info-Ratios on the previous year's TR^2s to investigate the predictive power of TR^2 on future alpha:

$$alpha_i\ (Info - Ratio_i) = \propto_i + \beta_i * TR_{i-1} + \varepsilon_i. \tag{6}$$

The above section can be summarised as follows: for every year and every fund, a regression is run and the R^2 is transformed into TR^2. Once all years have been analysed, alphas are regressed on the previous year's TR^2.

We then move to a more practical analysis. We repeat the procedure two times, once focusing on alpha and once focusing on Info-Ratio. Every year (every month in our modified approach), funds are ranked by TR^2 into quartiles and, within these quartiles, they are distributed into quartiles by alpha or Info-Ratio. We adopt two opposite ranking methods: a higher quartile implies a lower TR^2, while ranking by alpha (Info-Ratio) is conducted in a descending order, i.e. a higher rank signals a higher alpha (Info-Ratio). Accordingly, the best performing portfolio should be the 4th quartile portfolio out of the 4th quartile (we refer to it as portfolio 4:4). The funds in this portfolio will have the lowest TR^2s and, within the lowest TR^2s, the highest alphas (Info-Ratios). In other words, this portfolio will show the highest selectivity and the highest alphas (Info-Ratio) among all portfolios. This is the portfolio we track through the years. *Appendices 5* through *8* report summary statistics of all portfolios. *Appendices 5* and *6* refer to the original approach, while *Appendices 7* and *8* refer to the modified one. Additionally, *Appendices 5* and *7* refer to alpha-sorted portfolios, while *Appendices 6* and *8* refer to Info-Ratio based portfolios. The tables report minimum, maximum, mean and standard deviation of the average TR^2 and alpha (Info-Ratio) of each portfolio, where the average is taken over time and over the funds that make up each portfolio. Alphas (Info-Ratios) are reported in curly brackets. Results are fairly similar across all tables. As expected, TR^2 decreases within quartiles, while alpha (Info-Ratio) increases. This leads the 4:4 portfolios to present the lowest and highest values for the two variables. TR^2s are similar among approaches, while alphas naturally change because of different frequencies, which is annual in the original approach and daily in the modified one.

Once the portfolios are built, we calculate their returns based on the assumption that the funds are held for one year (or one month in the modified approach) and that portfolios are invested with equal weights in the funds. As discussed above, we assign a return of -20% to funds that cease to exist, disappear, or merge and, hence, do not report returns for the year. Finally, we run Carhart four-factor regressions for each portfolio's return. We run the same model as in *Equation 3*. We use monthly frequencies to have a larger number of data points, although we maintain the holding periods explained before. The regression fits the portfolios' excess returns to the Carhart four-factors, namely the AQR Market Index and the Swedish equivalents of the HML, SMB and UMD factors. Finally, we plot the cumulative performance of the 4:4 portfolios from both approaches, comparing them to the chosen benchmarks.

4.4 Statistical tests

We implement three steps to check for specific statistical characteristics of our data, namely the Durbin-Watson test, the Breusch-Pagan test and the Newey-West estimator. Each step serves a different purpose: the first checks for autocorrelation in the residuals of the dependent variable. The second tests for heteroscedasticity in the dependent variable and the third provides an estimate of the covariance matrix of parameters of a regression model. The steps are implemented on the Carhart four-factor regressions of portfolios' returns.

The Durbin-Watson test has a null hypothesis of no autocorrelation among the residual terms from a regression analysis. In other words, the test checks if the error terms of one period of data are linearly dependent on the error terms of the prior period.[9] The output of the function is the DW value and the corresponding p-value. In *Section (5)* below, p-values for the test are reported, as these are the criteria by which we reject, or fail to, the null hypothesis. It is important to test for autocorrelation as it, if present, lowers the efficiency of our regression model. Furthermore, it leads to incorrectly estimated standard errors and distorts t-statistics.

The Breusch-Pagan test has a null hypothesis of homoscedasticity, which implies constant variance of the errors in a regression model.[10] The output of the function is the BP value and the corresponding p-value. For the same reasoning as before, p-values for the test are reported in *Section (5)*. Hendricks et al. (1993) argue that the ranked portfolios may be heteroscedastic, because both the number and the identity of funds in a given portfolio change from quarter to quarter or from month to month. Heteroscedasticity means that the variance of the residual values and, hence, the variance of the dependent variable itself, is significantly different with

[9] In R, the test is implemented through the *dwtest* function from the *lmtest* package.
[10] In R, the test is implemented through the *Bptest* function from the *lmtest* package.

each value of the independent variable, i.e. it is not constant across all observations. As a consequence, our regression may be incorrectly specified.[11]

The Newey-West[12] estimator aims to overcome autocorrelation and heteroscedasticity in the error terms of the model. The output of the function is an adjusted covariance matrix of the parameters. We obtain Newey-West adjusted standard errors and, consequently, t-statistics for the parameters[13]. We note that in *Section (5),* reported t-statistics depict adjusted values. Original t-statistics, i.e. non-Newey-West adjusted, can be found in *Appendix 14*.

5. Results

Each predictor yields different results, both in terms of frequency, format and performance. The five-year estimation window of the original approach of the back testing predictor prevents us from comparing all predictors over a similar investment horizon. Additionally, the original approach of the momentum predictor provides results in quarterly frequency. Naturally, each predictor works in different ways and provides different end results, which are directly compared in *Section (6.1.)*.

This section is structured as follows: for each predictor, we comment on the results from the original methodology and then on the results from our modified approach. We report tables that summarise statistical results of each approach and plot the cumulative performance of each top portfolio against the benchmark indexes. The top portfolios represent each predictor's best performing portfolio, by construction.

5.1. Momentum Predictor

Following the steps described in the methodology section above, we obtain eight portfolios in each period for the momentum predictor. We run OLS regressions of each of these portfolios on the Carhart four-factor model and compare the performance of the supposedly best portfolio, i.e. the 8th octile portfolio, to the performance of our benchmarks: the AQR Market Index, the OMXS30 and the OMXSMCPI. Through preliminary tests, and consistently with the findings by Hendricks et al., we identify that an estimation window of one year yields the highest alphas, compared to using other estimation windows. Therefore, we focus on two specific approaches,

[11] While Hendricks et al. recommend using the White test (White 1980), a specification of the Breusch-Pagan methodology that has advantages in specific settings over the basic Breusch-Pagan approach, we implement the basic Breusch-Pagan test, mainly because the White test is part of a package not available for new versions of R anymore. However, different studies have demonstrated that the two tests may be equivalent.
[12] In R, the estimator is implemented through the *NeweyWest* function from the *sandwich* package.
[13] In R, this is done with the *coeftest* function from the *lmtest* package.

based on a one-year estimation window, using either quarterly or monthly data, introduced above as the original approach and the modified approach, respectively.

Tables 2 and *3* report summary statistics of the regressions for each portfolio we build. *Table 2* refers to the original approach, while *Table 3* applies to the modified approach. As such, the holding period for portfolios represented in *Table 2* is one quarter, whereas *Table 3* reports results for portfolios with a one-month holding period. Alphas in both tables represent annualised values.

Applying the momentum predictor to the Swedish mutual fund industry yields counter-intuitive results. Alphas from the Carhart four-factor regression are not increasing, but rather seem to be decreasing, with octiles. The 4^{th} octile is the portfolio with the highest alpha of 0.00148, while the 1^{st} octile portfolio has the second highest alpha (-0.00229). The 8^{th} octile portfolio, supposed to be the best performing, reports the third lowest alpha, at -0.01676. The 6^{th} octile portfolio is the only one that presents a statistically significant alpha at a 5% level, with a t-statistic of -1.7792. The significance values are adjusted with the Newey-West estimator. Original, non-adjusted values can be found in *Appendix 9*. Beta values with respect to the market index are, as expected, all very close to 1. This is due to the limited size of the Swedish stock market, which does not let fund managers vary much from the index. Although not reported, we note that market betas showed high significance levels. Exposure to other factors varies: the value factor is mostly negative, except for the 6^{th} octile portfolio. The size factor is positive for the lower octiles and negative for the last three, while momentum is overall positive, except for the 3^{rd} octile portfolio. It seems that the 8^{th} octile portfolio's return is affected by both the more long-term value premium and the more short-term momentum premium, while the size premium has a negative effect.

P-values for the Durbin-Watson test are very large with values around 0.5 or higher, meaning that we do not reject the null hypothesis of uncorrelated residual terms of the regression. Hence, autocorrelation seems to not be an issue in this case, meaning that the standard errors in period *t* are not correlated with standard errors of the preceding time period *t-1*. This is an important observation, as autocorrelated residuals lead to an underestimation of the sample variance, which in turn may lead to upwards skewed significance of regression coefficients. Similarly, we cannot reject the null hypothesis of the Breusch-Pagan test, stating homoscedasticity, for most of the portfolios, although the results are not as clear as they are for the Durbin-Watson test. Heteroscedastic data might lead to biased standard errors and distorted variance estimation. Hence, significance tests of the regression parameters would be skewed, potentially giving rise to false interpretations of our results. The 8^{th} octile portfolio presents a p-value of 0.0673, which we interpret as a weak sign of heteroscedasticity. As mentioned, these issues are accounted for and corrected using the Newey-West estimator.

Original Momentum Predictor

	Alpha	AQR Market Index	HML	SMB	UMD	Durbin Watson Test P-Value	Breusch-Pagan Test P-Value
Octile 1	-0.00229 (-0.1639)	0.9066	-0.1350	0.1435	0.0086	0.8527	0.1360
Octile 2	-0.02234 (-1.4266)	0.9384	-0.0330	0.0220	0.0189	0.6582	0.9760
Octile 3	-0.00507 (-0.3706)	0.9872	-0.0827	0.0028	0.0761	0.7323	0.5469
Octile 4	0.00148 (0.1400)	0.9844	-0.0621	0.0329	-0.0093	0.6043	0.4863
Octile 5	-0.01369 (-1.1065)	1.0047	-0.0993	0.1106	0.0016	0.6238	0.9163
Octile 6	-0.02746 (-1.7792)	1.0570	0.0503	-0.1240	0.0590	0.9864	0.1074
Octile 7	-0.00362 (-0.2395)	1.0530	-0.0204	-0.0783	0.0949	0.5474	0.4703
Octile 8	-0.01676 (-0.8797)	1.0596	0.0662	-0.0944	0.0974	0.4308	0.0673

Table 2 - Summary statistics from a Carhart four-factor regression of portfolios' excess returns – original approach of the momentum predictor
In every period, funds are sorted into octiles based on their performance during the previous four quarters and portfolios are built based on the octiles (a higher octile corresponds to a higher past performance). Each portfolio is held for one quarter, then the procedure is repeated. The first portfolio is built in 1Q2004 and the last in 3Q2016. We run a regression of each portfolio's returns on the four Carhart factors and report alphas and betas to each factor. The regression is based on quarterly data but alphas are annualised. Newey-West adjusted t-statistics are reported in brackets below each alpha. P-values from statistical tests, checking for autocorrelation and heteroscedasticity, are also reported.

Table 3 reports summary statistics from our modified approach. We change from a quarterly to a monthly analysis, which results in a higher turnover of funds within the portfolios. Please refer to *Section (4.1.)* for more details. The procedure seems to produce better results than the original approach, presented above. Annualised alphas from the Carhart four-factor regression are overall increasing with octiles. The 8th octile reports the highest alpha of 0.01041, the 4th octile reports the second highest alpha at 0.00821 while the 2nd octile reports the lowest value at -0.01626. In general, annualised alphas are higher with a shorter holding period, with an average of 0.0012 compared to -0.0112 in the original approach. On the other hand, none of the alphas are statistically significant at a 5% level. T-statistics are adjusted with the Newey-West estimator and unadjusted values can be found in *Appendix 10*. As in the original approach, betas for the market factor are fairly high at 0.90, although they are lower than in the previous case. This is probably due to the shorter time frame that we use in this approach, which might create more noise. Factor loadings behave as in the quarterly approach, apart from the SMB factor, which has a negative coefficient for all the octiles apart from the 1st octile portfolio. HML loadings are all negative, while UMD betas are positive. Unlike in the original approach, the 8th octile portfolio's return is positively affected by the momentum premium only, while the value premium and size premium have negative effects. This observation is consistent with our expectations, as the modified approach has a shorter evaluation period than the original approach. Hence, the momentum factor can be expected to have a bigger influence on portfolio returns than the longer-term HML and SMB factors.

Again, we use the Durbin-Watson test to check for autocorrelation, which does not seem to be an issue for the modified momentum approach. P-values from the test are all above 0.9, only the 8th portfolio shows weak signs of autocorrelation (p-value of 0.0672). However, we cannot reject the null hypothesis for any of the portfolios. The Breusch-Pagan test signals that the 1st octile portfolio is affected by heteroscedasticity. The test shows a p-value of 0.0150, which is lower than 5% and, hence, results in the rejection of the null hypothesis of homoscedasticity. This implies that standard errors from the regressions of this portfolio are likely to be biased, leading to distorted significance values of the regression parameters. Additionally, the 4th and the 8th octile portfolios show low p-values to the test, at 0.0998 and 0.0636, respectively. Overall, it seems that the modified approach increases heteroscedasticity in the data, with two more portfolios seemingly affected by it compared to the original methodology.

Modified Momentum Predictor

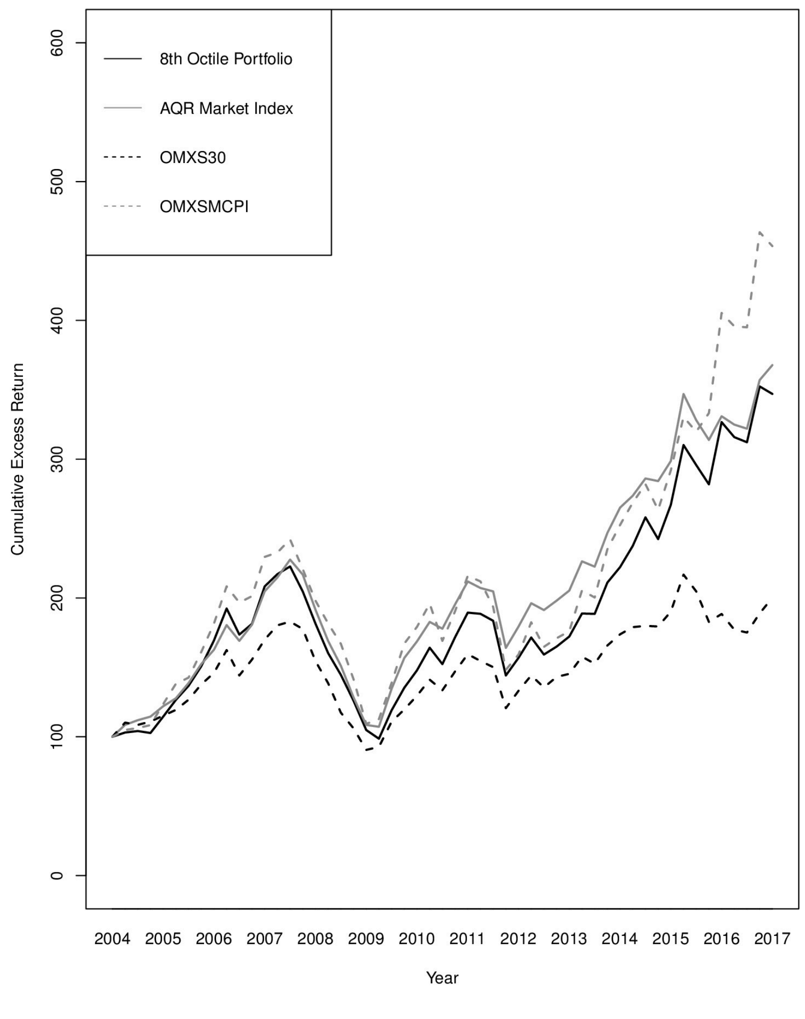
	Alpha	AQR Market Index	HML	SMB	UMD	Durbin Watson Test P-Value	Breusch-Pagan Test P-Value
Octile 1	0.00579	0.8809	-0.1047	0.0184	0.1036	0.9723	0.0150
	(0.3823)						
Octile 2	-0.01626	0.9185	-0.0277	-0.0785	0.0492	0.9985	0.8867
	(-1.1525)						
Octile 3	-0.00779	0.9561	-0.1157	-0.0519	0.1213	0.9994	0.4681
	(-0.6388)						
Octile 4	0.00821	0.9505	-0.0793	-0.0235	0.0681	0.9998	0.0998
	(0.7530)						
Octile 5	0.00289	0.9445	-0.0930	-0.0140	0.0909	0.9997	0.4627
	(0.2307)						
Octile 6	0.00251	0.9234	-0.0722	-0.0305	0.1037	0.9931	0.1877
	(0.1984)						
Octile 7	0.00355	0.9536	-0.0635	-0.0948	0.1603	0.9207	0.2484
	(0.2430)						
Octile 8	0.01041	0.9154	-0.0144	-0.1236	0.1372	0.0672	0.0636
	(0.5630)						

Table 3 - Summary statistics from a Carhart four-factor regression of portfolios' excess returns – modified approach of the momentum predictor

In every period, funds are sorted into octiles based on their performance during the previous twelve months and portfolios are built based on the octiles (a higher octile corresponds to a higher past performance). Each portfolio is held for one month, then the procedure is repeated. The first portfolio is built in January 2004 and the last in November 2016. We run a regression of each portfolio's returns on the four Carhart factors and report alphas and betas to each factor. The regression is based on monthly data but alphas are annualised. Newey-West adjusted t-statistics are reported in brackets below each alpha. P-values from statistical tests, checking for autocorrelation and heteroscedasticity, are also reported.

Figures 1 and *2* depict the plotted cumulative performance of the 8th octile portfolio from the two approaches over the 2004 – 2016 period. The two portfolios are compared to three benchmarks: the AQR Market Index, the OMXS30 and the OMXSMCPI. The frequencies of the graphs differ according to the construction of the portfolios, as they are quarterly for *Figure 1* (original approach) and monthly for *Figure 2* (modified approach).

In the original approach, the 8th octile portfolio presents a CAGR of 10.04%. This leads it to underperform the OMXSMCPI (12.33% CAGR), slightly underperform the AQR Market Index (10.56% CAGR) and outperform the main index OMXS30 (5.48%). The top portfolio from the modified approach slightly outperforms the original one, reporting a 10.31% CAGR compared to a 10.04% CAGR. The relative ranking against the benchmarks is the same for both portfolios.



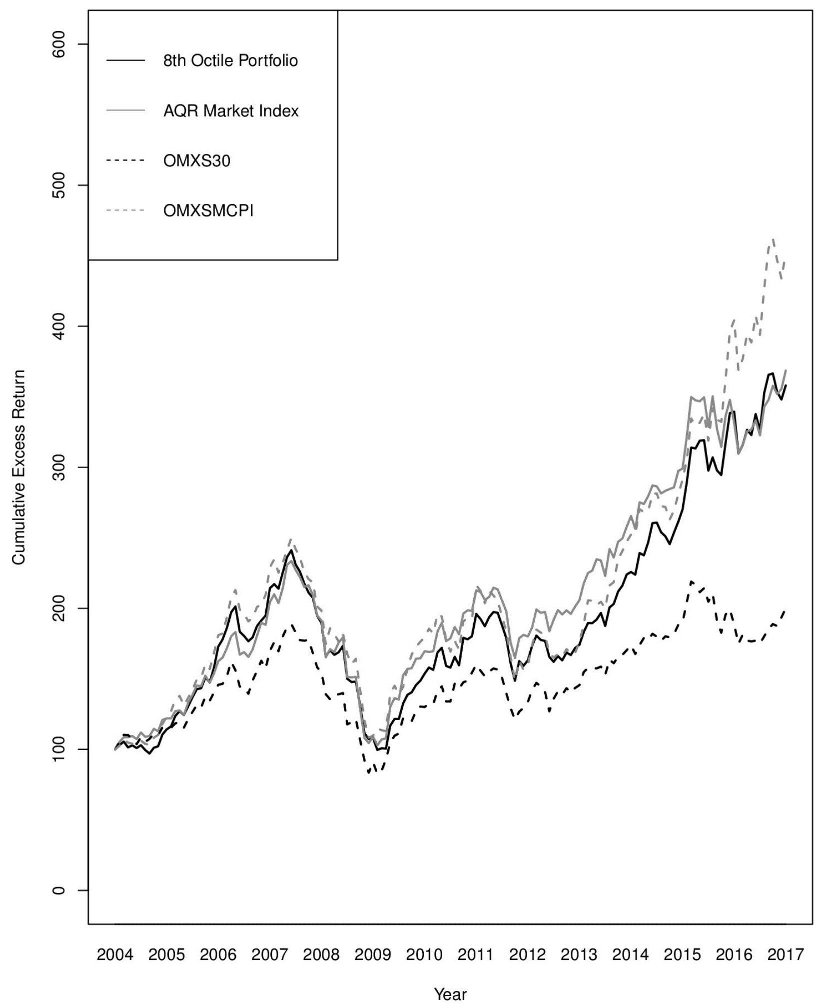


Figure 1 - Original approach 8th octile portfolio's cumulative performance vs. benchmark indexes

Benchmarks are the AQR Market Index, OMXS30 and OMXSMCPI. Data frequency is quarterly (original momentum predictor).

Figure 2 - Modified approach 8th octile portfolio's cumulative performance vs. benchmark indexes

Benchmarks are the AQR Market Index, OMXS30 and OMXSMCPI. Data frequency is monthly (modified momentum predictor).

5.2. Back Testing Predictor

We implement the back testing predictor following the original study of Mamaysky et al. (2007) as well as introducing a modified approach. Both approaches are explained in greater detail in *Section (4.2.)* above.

Table 4 reports summary statistics for the decile portfolios built following the methodology of Mamaysky et al. (2007), i.e. with an estimation period of five years and a holding period of one month. Portfolio excess returns are regressed on a Carhart four-factor model. It is worth noting that our decile order is "reversed", i.e. the 1st decile portfolio is supposed to be the best performing portfolio. A first look at the table reveals relatively meaningful results. Alphas are, by tendency, higher for low decile portfolios and decreasing with increasing decile number, ending up negative for the worst performing decile portfolios. However, the alpha of the 1st decile portfolio, which is 0.02008 annually, is only the fourth highest alpha we obtain. Oddly, alphas for the 2nd, 4th and 5th decile portfolios are higher than the ones reported by the top portfolio (1st octile portfolio). The 4th and 2nd decile portfolios both show substantially higher alphas than the 1st decile, with alphas of 0.03276 and 0.02927 annually, respectively. At the higher end of the decile portfolios, alphas behave more according to expectations, even though not perfectly. The 10th decile portfolio exhibits the second lowest alpha, while the 9th decile portfolio shows the lowest value by far (-0.01940 vs. -0.03441). None of the portfolios present alphas that are significant at a 5% level. We note that *Table 4* shows Newey-West adjusted results and report non-adjusted t-statistics in *Appendix 11*.

As can be expected, the market excess return is positively correlated to the portfolio excess returns. Betas range from 0.8166 to 0.9810 and are highly significant, although we do not report their t-statistics. The SMB factor shows mainly positive values between 0.03 and 0.23. The value size premium shows a negative relation to the portfolio excess returns, with similar magnitude as the UMD factor, with betas around -0.07. This is somewhat surprising as the value premium represents a rather long-term effect, while momentum naturally portrays short-term impact. The 1st decile portfolio's excess return is mostly explained by the market risk premium as well as the size premium (betas of 0.9810 and 0.2025, respectively). The value and the momentum factor show negative loadings, as for most other portfolios.

We use the Durbin-Watson test and the Breusch-Pagan test to check our results for autocorrelation and heteroscedasticity. P-values from both tests for each portfolio are reported in *Table 4*. The Durbin-Watson test suggests that the portfolios are not affected by autocorrelation, as all p-values are relatively large (p-values above 0.6) with the exception of the 6th decile portfolio (p-value of 0.0635). In fact, we cannot reject the null hypothesis of no autocorrelation, thus signalling that our data is sound. The 3rd decile portfolio registers a Breusch-Pagan p-value of 0.0002, which implies its residuals are heteroscedastic. P-values below the 5% hurdle lead to a rejection of the null hypothesis of homoscedasticity. Hence, we should be careful in interpreting the regression results for this specific portfolio, because heteroscedasticity implies biased standard errors and underestimation of sample variance, which in turn might lead to distorted significance tests of the regression coefficients. Nonetheless, all other portfolios seem not to be affected by heteroscedastic error terms (p-values between 0.1014 and 0.7712).

Original Back Testing Predictor

	Alpha	AQR Market Index	HML	SMB	UMD	Durbin Watson Test P-Value	Breusch-Pagan Test P-Value
Decile 1	0.02008	0.9810	-0.1450	0.2025	-0.0989	0.7183	0.1939
	(0.9328)						
Decile 2	0.02927	0.8880	-0.1379	0.1974	-0.1142	0.8312	0.1490
	(1.1269)						
Decile 3	0.01212	0.8166	-0.0790	0.0306	0.0012	0.6080	0.0002
	(0.4661)						
Decile 4	0.03276	0.9351	-0.1127	0.0889	-0.0530	0.7759	0.2500
	(1.5984)						
Decile 5	0.02335	0.9380	-0.0985	0.0813	-0.0947	0.7609	0.7189
	(1.0405)						
Decile 6	0.00002	0.9675	-0.0501	0.1441	-0.1265	0.9931	0.7712
	(0.0009)						
Decile 7	-0.00278	0.8854	0.0206	0.1150	-0.1188	0.0635	0.1014
	(-0.1232)						
Decile 8	-0.00682	0.9130	-0.1050	0.0598	-0.0557	0.8946	0.4120
	(-0.3782)						
Decile 9	-0.03441	0.9805	-0.1123	0.0645	-0.0143	0.9773	0.6103
	(-1.3298)						
Decile 10	-0.01940	0.9601	-0.0510	0.2251	-0.1220	0.9967	0.1039
	(-0.9505)						

Table 4 - Summary statistics from a Carhart four-factor regression of portfolios' excess returns – original approach of the back testing predictor
Every month, funds are filtered through a two-step back test. The first step compares the estimated alpha to the next realised alpha. The estimated alpha is from a Carhart four-factor regression of each fund's excess returns, based on 60 months' worth of data. The alpha from this regression is compared to the realised excess return in month 61 (fund's return in excess of market return) and, if the sign of the two is the same, the fund stays in the active pool. The next step involves a second Carhart four-factor regression, with a 1-month forward shift relative to the first step. If estimates of alpha and beta (relative to the Market factor) fall within a specified interval, the fund remains in the active pool. Funds in the active pool are ranked by alpha into deciles. We employ a reversed ranking which results in the best-performing funds populating the 1st decile. Equally-weighted portfolios are built based on the deciles and are held for one month, after which the procedure is repeated. Due to the five-year estimation window, the first decile-based portfolios are built in January 2008, while the last ones are built in December 2016. Alphas and betas are obtained from a Carhart four-factor regression of the portfolios' returns. Alphas are annualised as the regression yields monthly alphas. Newey-West adjusted t-statistics are presented in parentheses below the according alpha. P-values from statistical tests, checking for autocorrelation and heteroscedasticity, are also reported.

Table 5 presents summary statistics for all decile portfolios following our modified back testing approach, which uses an estimation window of one year, while keeping the holding period at one month. T-statistics are adjusted using the Newey-West methodology. Unlike the results for the original approach above, several portfolios show signs of heteroscedasticity, which makes the Newey-West adjustment necessary (refer to *Appendix 12* for unadjusted results). The alphas we obtain from the modified approach distribute more closely around zero compared to the original approach. However, in this case, the 1st decile portfolio exhibits the highest alpha by far, with 0.04236, which is also significant at a 5% level. The 10th decile portfolio, supposed to be the worst, presents an annualised alpha of 0.00193. Almost all alphas are positive, which is questionable, with the exception being the 9th decile portfolio with an annualised alpha of -0.00487. While the 1st decile portfolio's alpha is statistically significant, we note it is the only one that is. Our adjusted approach produces better results than the original methodology. The average alpha for the original study amount to 0.00542, which is well below the average alpha of 0.01220 that our modified approach yields.

Overall, alpha decreases inconsistently with increasing decile number. Betas of the market excess return range from 0.7092 to 0.9739. The lower market betas compared to the original approach might arise from the switch to daily data. We discuss this potential issue and effects for our regressions in *Section (6.3.)*. UMD factor loadings slightly increase, ranging from -0.0904 to 0.0270. The betas for value and size factor are very similar to the those from the original approach, which is to be expected, as the holding period does not change in our modification. The 1st decile portfolio's excess return is strongly influenced by the market risk premium, as the beta is at 0.9739. The SMB factor coefficient is lower than in the original approach, with a value of 0.1240. The HML and the UMD factor coefficients remain negative, although they are higher than in the original approach. Autocorrelation does not seem to be an issue in the modified approach as the Durbin-Watson test never exhibits a p-value below 5%. Hence, the null hypothesis of no autocorrelation cannot be rejected in any instance. The Breusch-Pagan test, on the other hand, presents significant results for all but one portfolio (6th decile), which suggests that the modified approach creates heteroscedasticity in the error terms. This observation outlines the need for Newey-West adjustments of the variance-covariance matrix to overcome the heteroscedasticity in the residual terms and make the regression output interpretable.

Modified Back Testing Predictor

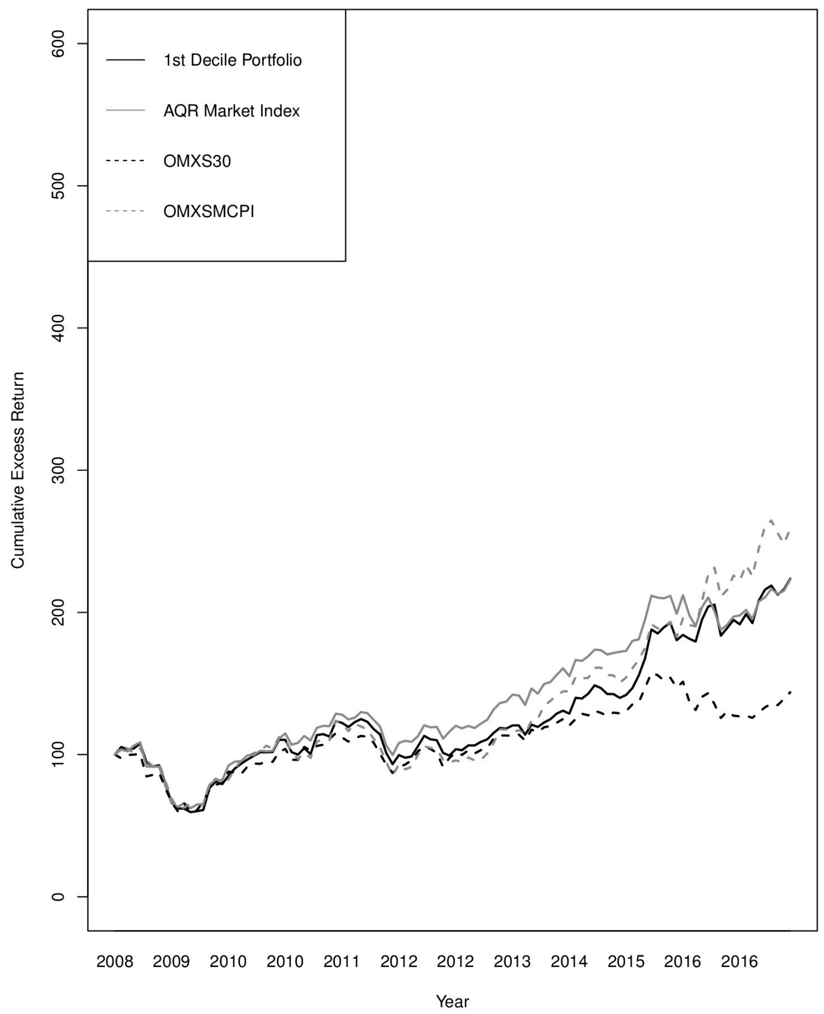
	Alpha	AQR Market Index	HML	SMB	UMD	Durbin Watson Test P-Value	Breusch-Pagan Test P-Value
Decile 1	0.04236	0.9739	-0.0809	0.1240	-0.0767	0.7592	0.0498
	(2.1591)						
Decile 2	0.01004	0.8803	-0.0031	0.0552	-0.0201	0.1093	0.0060
	(0.4808)						
Decile 3	0.02555	0.8649	-0.0147	0.1502	-0.0904	0.7604	0.0114
	(1.3265)						
Decile 4	0.00651	0.8288	-0.1300	0.1234	-0.0015	0.9851	0.0249
	(0.2861)						
Decile 5	0.01040	0.8375	-0.0572	0.1819	-0.0853	0.4644	0.0226
	(0.4913)						
Decile 6	0.00479	0.9001	-0.0967	0.1391	-0.0411	0.9882	0.2662
	(0.2804)						
Decile 7	0.01335	0.8644	-0.1017	0.1081	0.0217	0.9619	0.0478
	(0.5383)						
Decile 8	0.01199	0.8121	-0.0777	0.2338	-0.0485	0.7909	0.0472
	(0.6347)						
Decile 9	-0.00487	0.8717	-0.1267	0.1316	-0.0113	0.9536	0.0056
	(-0.2920)						
Decile 10	0.00193	0.7092	-0.1233	0.2678	0.0270	0.2112	0.0061
	(0.0577)						

Table 5 - Summary statistics from a Carhart four-factor regression of portfolios' excess returns – modified approach of the back testing predictor

Every month, funds are filtered through a two-step back test. The first step compares the estimated alpha to the next realised alpha. The estimated alpha is from a Carhart four-factor regression of each fund's excess returns, based on 12 months' worth of daily data. The alpha from this regression is compared to the realised excess return in month 61 (fund's return in excess of market return) and, if the sign of the two is the same, the fund stays in the active pool. The next step involves a second Carhart four-factor regression, with a 1-month forward shift relative to the first step. If estimates of alpha and beta (relative to the Market factor) fall within a specified interval, the fund remains in the active pool. Funds in the active pool are ranked by alpha into deciles. We employ a reversed ranking which results in the best performing funds populating the 1st decile. Equally-weighted portfolios are built based on the deciles and are held for one month, after which the procedure is repeated. Due to the one-year estimation window, the first decile-based portfolios are built in January 2004, while the last ones are built in December 2016. Alphas and betas are obtained from a Carhart four-factor regression of the portfolios' returns. Alphas are annualised as the regression yields monthly alphas. Newey-West adjusted t-statistics are presented in parentheses below the according alpha. P-values from statistical tests, checking for autocorrelation and heteroscedasticity, are also reported.

Figure 3 depicts the cumulative performance of the 1st decile portfolio following the original approach of Mamaysky et al. (2007). We plot the cumulative performance against three benchmarks (OMXS30, OMXSMCPI and the AQR Market Index, the last is the market index used in the four-factor regression). The portfolio and the AQR Market Index have the same overall performance, although the index outperforms the portfolio throughout most of the sample. Between 2008 and 2016, the AQR Market Index shows a CAGR of approximately 9.33%, while the 1st decile portfolio performs slightly better, with a CAGR of 9.36%. Both performances are somewhat mediocre compared to the OMXSMCPI, which exhibits a CAGR of approximately 11.15% over the same time horizon. Hence, as is the case with the momentum predictor, the portfolio underperforms the OMXSMCPI, while outperforming the OMXS30. The main index, as a matter of fact, has a CAGR of 4.12%.

Figure 4 shows the cumulative performance of the 1st decile portfolio according to our modified approach. Again, we plot the three benchmarks OMXS30, OMXSMCPI and the AQR Market Index. Judging from the graph, the modified back testing approach seems to perform very well. Not only does it substantially outperform the 1st decile portfolio of the original approach, it also outperforms all three benchmarks. The CAGR of the 1st decile portfolio comes in at 12.87%, substantially higher than the one from the original approach. The portfolio outperforms all three indexes, i.e. the OMXSMCPI (12.33% CAGR), the AQR Market Index (10.56% CAGR) and the OMXS30 (5.48% CAGR).



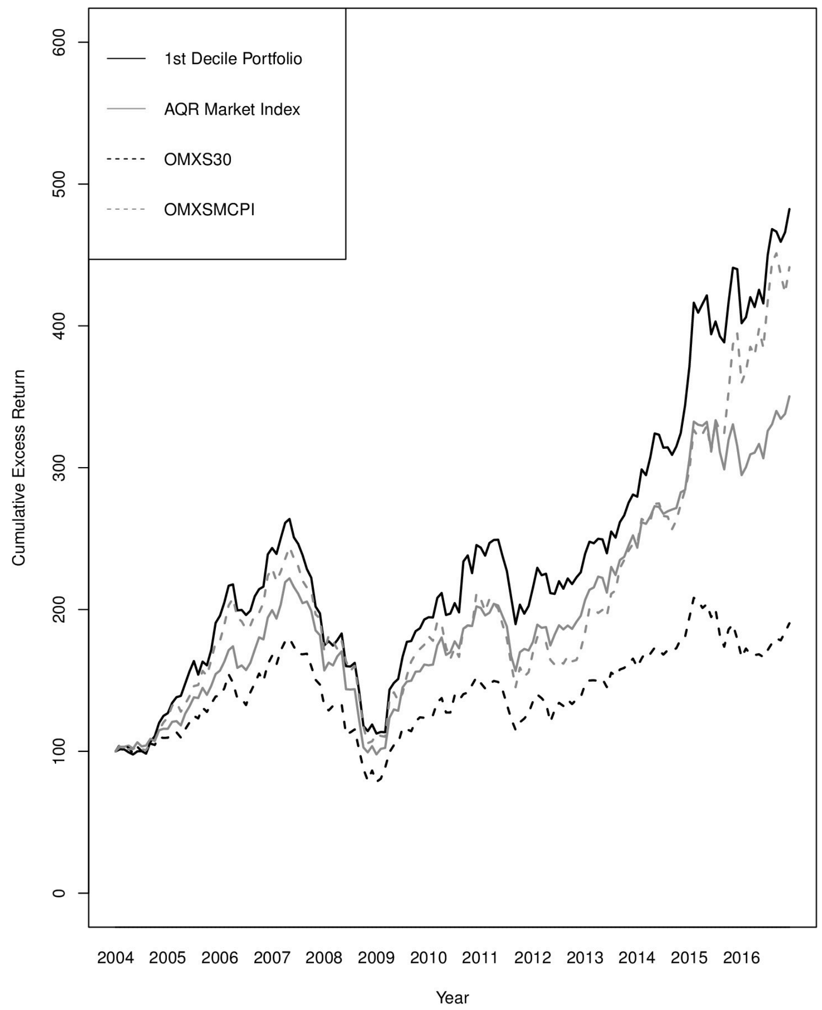


Figure 3 - Original approach 1st decile portfolio's performance vs. benchmark indexes

Benchmarks are the AQR Market Index, OMXS30 and OMXSMCPI. Data frequency is monthly, starting in January 2008 (original back test predictor).

Figure 4 – Modified approach 1st decile portfolio's performance vs. benchmark indexes

Benchmarks are the AQR Market Index, OMXS30 and OMXSMCPI. Data frequency is monthly, starting in January 2004 (modified back test predictor).

5.3. Selectivity Predictor

We follow Amihud and Goyenko (2013) and investigate the predictive power of TR^2, a logical transformation of R^2 of future alpha. As illustrated in the methodology section, we first run a regression of each fund's return on the Carhart four-factors. As we want to investigate the predictive power of TR^2, we regress alpha (Info-Ratio) on the previous year's $TR^{2.}$ *Appendix 14* reports the summary statistics of those regressions. We obtain stronger and more significant results in our modified approach. Additionally, TR^2 seems to have a stronger, negative predictive power on alpha than on Info-Ratio. Consistently with Amihud and Goyenko's hypothesis and findings, all coefficients are negative. TR^2 presents a coefficient of -0.05722 on future alpha in the original approach (-0.08508 in the modified approach) while the values are -0.03727 and -0.05614, respectively, on future Info-Ratio. All coefficients are very significant, shown by high t-statistics, which are adjusted with the Newey-West estimator. We note that not adjusting does not present lower significance. Non-adjusted values are presented in the same table. The R^2s of the regressions we report are fairly low, particularly in the original approach, with values around 6% in the original approach and approximately 16% in the modified one. We believe the low R^2 is due to two main factors: the number of observations, which is clearly higher in the modified approach, and the lack of control variables, which are numerous in the original paper by Amihud and Goyenko (2013), but that we are missing entirely.

The Durbin-Watson test seems to detect autocorrelation in the residuals. In fact, all p-values show significance, regardless of methodology (p-values range between 0.00006 and 2.2e-16). Consequently, this leads us to believe the data is autocorrelated. The Breusch-Pagan test also shows significance overall, except for the regression of alpha on TR^2 in the original approach (p-value of 0.7531). This is a strong indicator of heteroscedasticity in the regression residuals. These issues are corrected for with the Newey-West estimator. Following the adjustment, we find that t-statistics decrease, although they remain highly significant, for both the alpha and the beta of the regression.

Table 6 reports summary statistics of a Carhart four-factor regression of each portfolio's excess returns. For each approach and sorting method (alpha or Info-Ratio) we build sixteen portfolios, as we rank funds by TR^2 in quartiles and again, within those, by alpha or Info-Ratio. A higher TR^2 quartile corresponds to a lower TR^2, and so to a higher selectivity and, hence, a higher expected performance. A higher alpha (Info-Ratio) quartile implies a higher alpha (Info-Ratio) and should correspond to higher performance and, consequently, to a higher alpha. As such, portfolio 4:4 is supposed to be the best performing portfolio.

The alpha-sorted portfolios from the original approach follow the predicted pattern only partially. While the highest TR^2 quartile does correspond to the highest alphas, this is not always true for the other quartiles. The same can be said about the alpha quartiles. In particular, the worst performing portfolio is 3:3, with an annual alpha of -0.00615. The best performing portfolio is 4:3, with an annual alpha of 0.04600, followed by 4:4, with an annual alpha of 0.04406. Significance is indeed growing in TR^2 quartiles, with all alphas of portfolios in the 4th quartile significant. The expected pattern is more visible in the modified approach, although the 2nd TR^2 quartile performs worse than the 1st. The pattern is more pronounced for alpha quartiles. The best performing portfolio is 4:2, with an annual alpha of 0.04617. Portfolio 4:4 ranks 3rd with an annual alpha of 0.04141. The worst performing portfolio is 2:1, with an annual alpha of -0.00835. Again, we find that the 4th TR^2 portfolios present significant alphas. These four, with the addition of portfolio 3:4, are the only portfolios with alphas significant at, at least, a 5% level. Overall, we note that the methodologies seem to work better for higher quartiles. At the lower end of the portfolios, supposedly containing the worst performing funds, the construction of the portfolios does not provide the expected results. Alphas are positive in almost all first quartile TR^2 portfolios. Out of 64 constructed portfolios, only seventeen exhibit negative alphas, which is somewhat surprising.

Among the Info-Ratio-sorted portfolios, the expected patterns appear more evident than above: portfolio 4:4 is the highest performing with an annual alpha of 0.05638, followed by 4:3 (annual alpha of 0.04554). However, the worst performing portfolio is 2:3, with an annual alpha of -0.00789. Portfolios containing better performing funds tend to be more significant. Other portfolios in the 4th R^2 quartile are all significant, however these are the only significant portfolios. In the modified approach, the 4th TR^2 quartile contains the highest performing portfolios, with portfolio 4:4 reporting the highest annual alpha of 0.06251. The worst performing portfolio is 2:1, with an annualised alpha of -0.00587. As for alpha-sorted portfolios, the 4th TR^2 quartile portfolios exhibit statistically significant alphas, and so does portfolio 3:4, supporting our impression that the methodology works better for high performing portfolios. We note that the t-statistics reported in *Table 6* are adjusted with the Newey-West estimator to account for autocorrelation and heteroscedasticity in the residuals of the regression model, described below. The non-adjusted values for the t-statistics can be found in *Appendix 13*.

For simplicity, we only describe the factor loadings of the 4:4 portfolios for each of the four approaches. Generally, the factor loadings resemble the ones we observe for the back testing predictor. For all four selectivity approaches, the coefficient of the market risk premium lies between 0.9157 and 0.9954. Further, the SMB factor has a positive relation to the 4:4 portfolios' excess returns, with reported betas between 0.1949 and 0.2296. The other two factors, i.e. the value and the momentum premium, exhibit negative coefficients around -0.1. Notably, all factors show high significance, with the exception of the value premium.

Appendices 15 and *16* report the p-values of the Durbin-Watson and Breusch-Pagan tests for each portfolio built following the selectivity predictor's methodology. On the one hand, autocorrelation does not seem to be an issue for the selectivity methodology. *Appendix 15* shows that the p-values are relatively large, lying between 0.4341 and 1.0000. Hence, the null hypothesis of uncorrelated residual terms of the four-factor regressions. On the other hand, heteroscedasticity is present in several portfolios throughout the different approaches. *Appendix 16* shows that each approach includes at least four portfolios that suffer from heteroscedasticity, as their p-values are small enough to reject the null hypotheses of homoscedastic error terms. In the original approach, two portfolios test positive to the Breusch-Pagan test, regardless of the sorting on alpha or Info-Ratio: portfolio 4:1 and portfolio 2:4. Two other alpha-sorted portfolios (1:3 and 3:4) and three Info-Ratio sorted portfolios (4:2, 3:3 and 4:4) have significant p-values from the Breusch-Pagan test. Turning to our modified approach, we see more consistent results: three portfolios (1:1, 3:4 and 4:4) test positive to the Breusch-Pagan test in both alpha and Info-Ratio ratio sorting. Portfolios 2:1 and 3:2 test positive when sorted on alpha and portfolio 4:3 when sorted based on Info-Ratio. As a consequence, these portfolios are all exposed to skewed variance estimations and distorted significance tests of the regression parameters, which most likely lead to misinterpretation of the results.

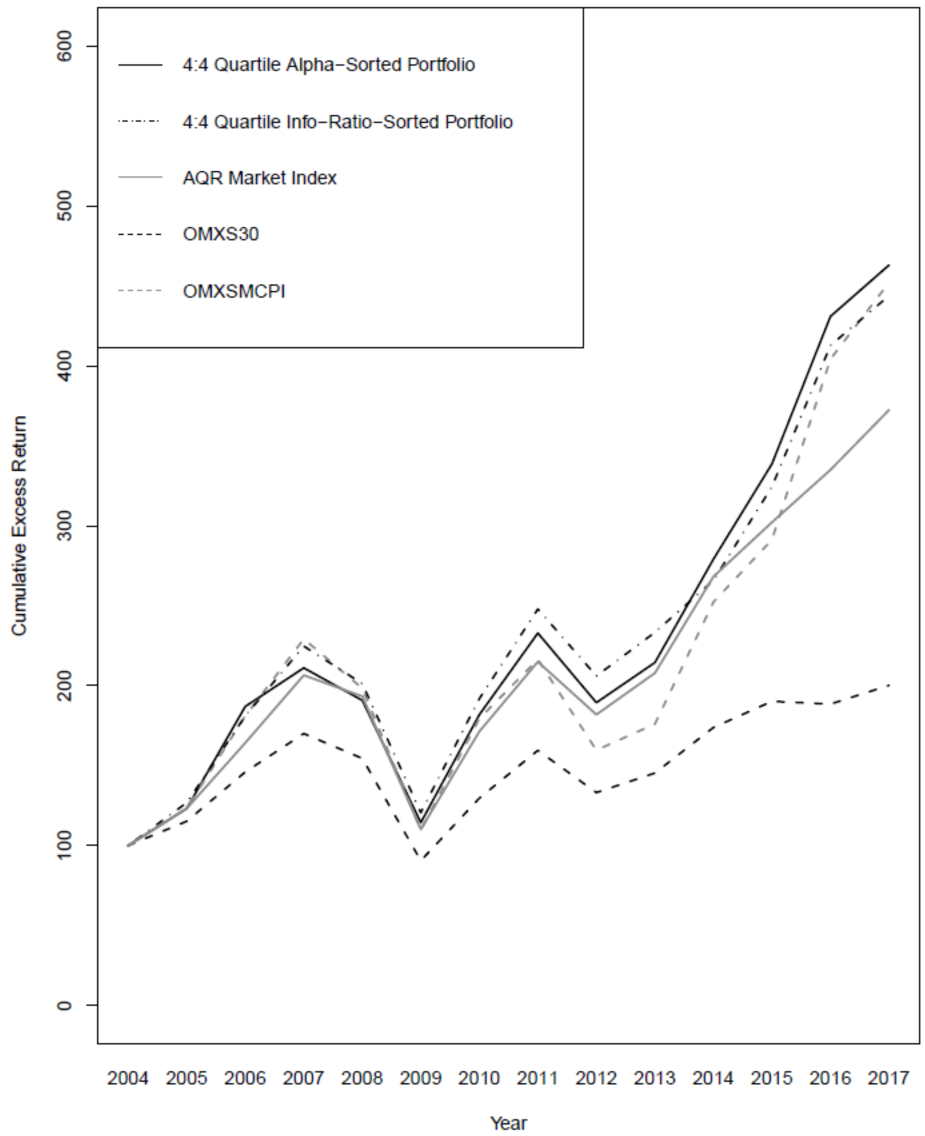
Our results suggest that actively managed mutual funds with a high selectivity, expressed by a low TR^2, earn higher excess returns than mutual funds that express a low selectivity. This finding directly feeds the notion that exceptionally skilled fund managers can produce positive alpha, at the very least over a certain period. While differences in TR^2 certainly arise by varying levels of fund selectivity, market timing most likely also has a bearing effect on TR^2. Naturally, this aspect can only be observed in actively managed funds. However, as Amihud and Goyenko (2013) reason that the effect is most likely limited in our study, due to the short estimation period of one year. Arguably, TR^2 is possibly closely linked to fund size. As smaller funds are probably less restricted in stock selection, meaning that they follow niche strategies, they are likely to have lower TR^2. Larger funds that have stricter mandates usually cover a broad set of industries and, hence, are more likely to have higher TR^2. This makes it somewhat difficult to derive a clear causality.

Selectivity Predictor

	Original Approach					**Modified Approach**			
	Alpha$_{t-1}$					Alpha$_{t-1}$			
R^2_{t-1}	Low	2	3	High	R^2_{t-1}	Low	2	3	High
Low	0.00051	0.01345	0.00982	-0.00171	Low	0.00516	0.00768	0.00215	0.01285
	(0.045)	(1.053)	(0.893)	(-0.153)		(0.444)	(0.677)	(0.186)	(0.971)
2	0.00402	0.00111	-0.00416	-0.00163	2	-0.00835	-0.00211	-0.00085	-0.00564
	(0.299)	(0.088)	(-0.385)	(-0.134)		(-0.699)	(-0.161)	(-0.075)	(-0.478)
3	0.02232	-0.00025	-0.00615	0.02808	3	0.00610	0.00780	0.00556	0.02545
	(1.538)	(-0.020)	(-0.495)	(1.856)		(0.454)	(0.619)	(0.404)	(1.849)
High	0.03210	0.03413	0.04600	0.04406	High	0.04011	0.04617	0.04279	0.04141
	(2.113)	(1.902)	(2.567)	(2.262)		(2.761)	(2.577)	(2.649)	(1.897)
	Info-Ratio$_{t-1}$					Info-Ratio$_{t-1}$			
R^2_{t-1}	Low	2	3	High	R^2_{t-1}	Low	2	3	High
Low	-0.00047	0.00713	0.01695	-0.00350	1	0.01115	0.00279	0.00079	0.01471
	(-0.043)	(0.584)	(1.315)	(-0.285)		(0.913)	(0.238)	(0.072)	(1.154)
2	0.00915	0.00015	-0.00789	-0.00141	2	-0.00587	-0.00027	-0.00541	-0.00427
	(0.649)	(0.013)	(-0.668)	(-0.125)		(-0.473)	(-0.021)	(-0.443)	(-0.363)
3	0.02016	-0.00655	0.00604	0.02278	3	0.00461	0.00470	0.00756	0.02835
	(1.369)	(-0.508)	(0.508)	(1.552)		(0.330)	(0.366)	(0.553)	(1.990)
High	0.02648	0.02962	0.04554	0.05638	4	0.03754	0.03000	0.04480	0.06251
	(1.703)	(1.640)	(2.448)	(3.019)		(2.278)	(1.732)	(2.703)	(3.017)

Table 6 - Summary statistics from a Carhart four-factor regression of portfolios' excess returns – selectivity predictor
In every period (every year in the original approach, every month in the modified approach), each fund's daily excess returns are regressed on the four Carhart factors. The regression spans over one year of historical data. The R^2 of the regression is retained and transformed into TR^2, where $TR^2 = \log(\sqrt{(R^2/(1-\sqrt{(R^2)}))})$. Funds are ranked into quartiles by TR^2 and, within these quartiles, by alpha (Info-Ratio). TR^2s are reversely ordered, so that quartile 4 includes the funds with the lowest TR^2s, which should perform better. Each portfolio is held for one year (original approach) or one month (modified approach). When the holding period expires, the procedure is repeated. We compute excess returns for the portfolios and regress them on the four Carhart factors. We note that in the original approach we calculate monthly returns to use in the regression, to increase the number of data points. We report annualised alphas for each portfolio and, below them, Newey-West adjusted t-statistics.

Figures 5 and *6* plot the cumulative returns of the 4:4 portfolios from our analyses versus the benchmarks OMXS30, OMXSMCPI and the AQR Market Index. In particular, *Figure 5* portrays the original approach, while *Figure 6* represents our modified approach. In *Figure 5*, the alpha-sorted portfolio is the best performer, with a CAGR of 12.51% over the 2004 – 2016 period. The OMXSMCPI comes second (12.33% CAGR), followed by the Info-Ratio-sorted portfolio (12.15% CAGR). The OMXS30 is, as always, the worst performing, with a CAGR of 5.48%. We note that the original approach assumes a holding period of one year, which makes the graph somewhat less detailed. In *Figure 6*, the Info-Ratio-sorted is the clear winner, with a CAGR of 14.40%. The alpha-sorted portfolio follows with a CAGR of 12.82%. In the second graph, the holding period is one month, which adds insights to the plot.



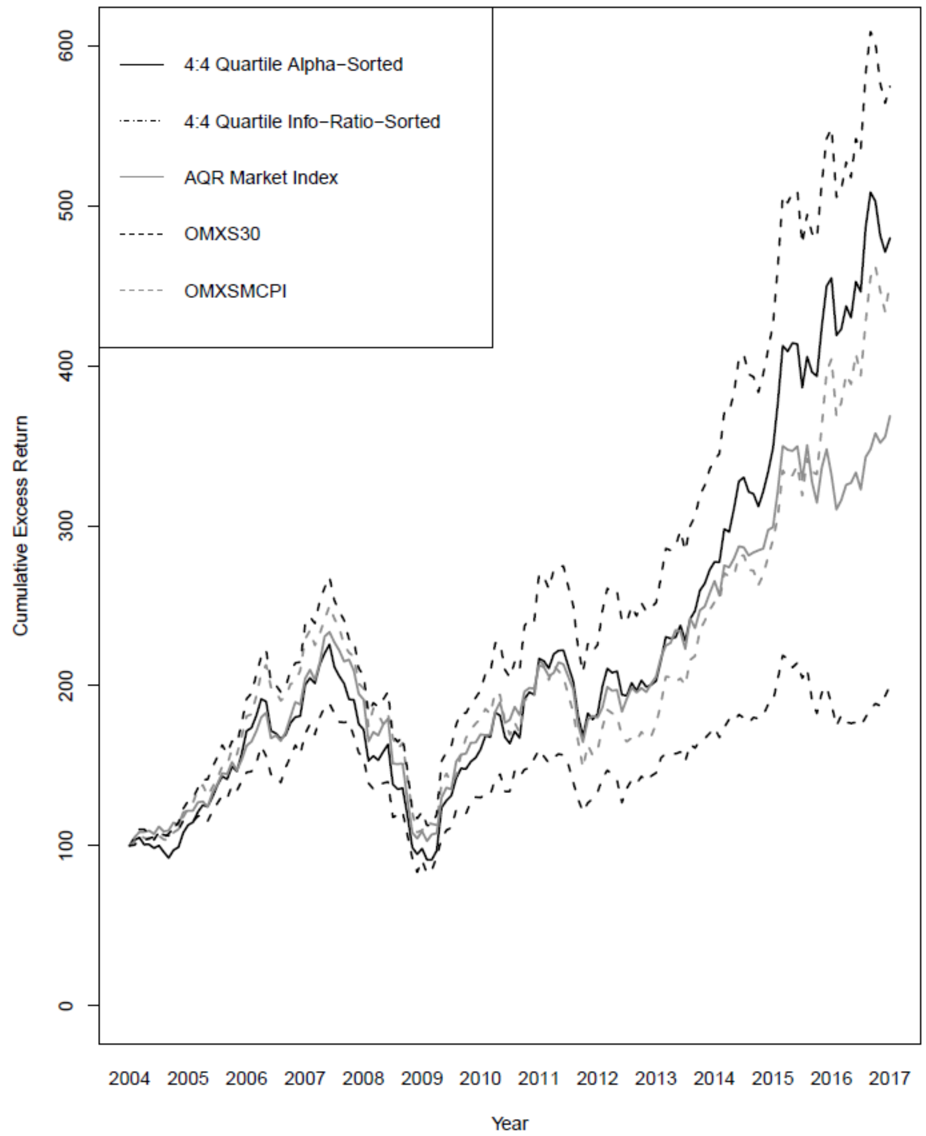


Figure 5 - Original approach 4:4 portfolios' performance vs. benchmark indexes

Both alpha- and Info-Ratio-sorted portfolios are plotted. Benchmarks are the AQR Market Index, OMXS30 and OMXSMCPI. Data frequency is annual, starting in 2004 (original selectivity predictor).

Figure 6 - Modified approach 4:4 portfolios' performance vs. benchmark indexes

Both alpha- and Info-Ratio-sorted portfolios are plotted. Benchmarks are the AQR Market Index, OMXS30 and OMXSMCPI. Data frequency is monthly, starting in January 2004 (modified selectivity predictor).

6. Critical Review

In the following sections, we analyse the three predictors from an economical point of view, comparing them on the basis of the cumulative performance that an investor would realise by investing in each of them. Additionally, we briefly compare each predictor's results from the Swedish context to those from its original paper, keeping in mind the differences in the base data. Finally, we reflect on general and predictor-specific considerations, and highlight the opportunities for future research.

6.1. Discussion of Results

The results presented above show that the three mutual fund performance predictors, namely the momentum predictor, back testing predictor and the selectivity predictor, exhibit different levels of predictive power. The momentum predictor is somewhat disappointing, as we do not achieve our main goal of producing meaningful economic gains from its implementation. The original approach for the back testing predictor yields similar results. However, the modified back testing predictor and the selectivity predictor produce significant alphas and outperform all three benchmark indexes OMXS30, OMXSMCPI and the AQR Market Index. In terms of economic outcome, the selectivity predictor shows the most impressive performance, particularly the Info-Ratio-sorted approach. *Table 7* reports the annualised alpha registered by each predictor's top portfolio (8th octile portfolios for the momentum predictor, 1st decile portfolios for the back testing predictor and 4:4 portfolios for the selectivity predictor) as well as the largest annualised alpha for each predictor. Further, we report CAGRs for the portfolios in the first column and their total holding periods. We conclude with a column showing the total gain from an initial investment of SEK100 in each portfolio from the first column. As mentioned above, we distinguish between the original methodologies and our own alterations, which mainly consist of changes in the estimation and holding periods. Some of the alterations we introduce are due to our limited dataset, which is much narrower than the ones available for the US market.

We note that our modified approaches lead to a better performance compared to the original ones in each case. The momentum predictor is the worst performing predictor amongst the three we analyse. In the original approach, the 8th octile portfolio reports a CAGR of 10.04%, backed by an alpha of -0.01676, which is not significant. This is not the highest alpha out of all momentum based portfolios, which report alphas as high as 0.00148. We conclude that our findings do not back those of Hendricks et al. (1993), which might be due to the different markets. Additionally, none of the alphas are significant. The momentum predictor is also the worst performing when looking at modified approaches, with a CAGR of 10.31%. In this case, the 8th octile portfolio is the best performing amongst its peers, with an annual alpha of 0.01041, which is still not significant. We define as peers the other portfolios built with the same methodology.

The different approaches are not comparable for the back testing predictor, as the approaches have different investment windows, due to the five-year estimation window of the original approach. The lower CAGR of the original approach (9.36% compared to 12.87%) can be explained by the weaker performance of the equity market during its investment window which, by coincidence, starts in 2008, at the height of the recent financial crisis. The 1st decile portfolio from the original approach is not the best performing portfolio, contrary to what we would expect. The portfolio reports an annual non-significant alpha of 0.02008 while the highest alpha for its peers – in this instance, peers refers to the 2nd through 10th decile portfolios built with the original back testing approach – is an almost significant 0.03276. The modified back testing predictor performs much better, with a CAGR of 12.87% and a statistically significant alpha of 0.04236. The 1st decile portfolio is the best performing amongst the other decile portfolio, and records the second highest performance of all "modified" predictors. As mentioned, we explain the better performance of the modified back testing predictor with the simple fact that it allows the predictor to start earlier, and not in the middle of the recent financial crisis. Additionally, a lower estimation window in this approach (one year compared to the five-year estimation window required by the original approach) may give better results regarding the alpha filtering.

The selectivity portfolio is clearly the best performing of the three, regardless of the sorting on alpha or Info-Ratio. For this predictor, we expect that the 4:4 portfolio of every approach would be the best performing. In the original approach, the alpha-sorted portfolio outperforms the Info-Ratio based one, with a CAGR of 12.51% compared to 12.15%. Strikingly, it reports a lower alpha (0.04406 compared to 0.05638, both significant). The lower alpha implies a higher loading on the common risk factors, which can explain the portfolio's higher performance. We conclude that the alpha-sorted predictor has a higher exposure to the risk factors, which seems to pay off. In particular, the alpha-based portfolio has a higher loading on the size factor, where it reports a beta of 0.2148 compared to the 0.1977 beta of the Info-Ratio-based portfolio.

In general, the ranking process does not seem to work well across selectivity-based approaches. In the case of alpha-based sorting, the 4:4 portfolios are outperformed by lower rank portfolios. However, both Info-Ratio-based approaches exhibit the highest alphas in portfolios 4:4.This approach only satisfies our expectations that a higher Info-Ratio and a lower TR^2 lead to a higher performance by tendency, as the pattern does not hold throughout all portfolios. Among the different approaches of the selectivity predictor, the modified approaches both outperform their original counterparts. The modified Info-Ratio sorted portfolio exhibits a CAGR of 14.40%, and is the best performing portfolio across our study. This is backed by an annualised alpha of 0.06251, which is statistically significant and shows that this performance is not reached by only loading on the common risk factors. The alpha-based portfolio reports the third highest CAGR of 12.82%, with a statistically significant alpha of 0.04141 annually.

We conclude that our modifications are beneficial to the performance of the predictors. The back testing predictor shows the highest improvement. However, this might mainly be because the modifications enable it to start working before the recent financial crisis. CAGRs for all predictors break the double-digit threshold, except for the original back testing predictor, and perform at least as the general Swedish equity market measured by the AQR Market Index. The back testing predictor also reports an increase in significance for its top portfolio's alpha. Our results, as depicted in the summary *Table 7*, include four top portfolios per approach, i.e. four original portfolios and four modified ones. The selectivity predictor yields two portfolios, one based on alpha sorting and one based on Info-Ratio. In the modified approach, three out of four portfolios, compared to two out of four in the original approach, show a statistically significant alpha. The modified approaches also seem to work better regarding the ranking process. According to our expectations, each of the portfolios we analyse ought to be the best performing among its peers. This is only true for one portfolio built with the respective original approach (Info-Ratio based selectivity predictor), while it is true for three out of four portfolios built following our modified approaches.

It is worth mentioning that all alphas depict values gross of fees and expenses, hence, those values are likely to be considerably lower. However, we note that these are the same for each predictor and, thus, should not change the direct comparison of the different approaches. We conclude that the selectivity predictor appears to be the best performing, and that R^2 is a well-functioning and easily implementable indicator of mutual fund selectivity. Moreover, it proves to be a well-functioning predictor of near-term future fund performance.

Original Approaches

Predictor	Top Portfolio	Largest Alpha	Top Portfolio's CAGR	Evaluation Period	SEK100 invested in 2004 and held until 2016
Momentum	-0.01676 (-0.8797)	0.00148 (0.1400)	10.04%	2004-2016	347.02
Back Testing	0.02008 (0.9328)	0.03276 (1.5984)	9.36%	2008-2016	223.65[a]
Selectivity Alpha-Based	0.04406 (2.2616)	0.04600 (2.5668)	12.51%	2004-2016	463.00
Selectivity Info-Ratio-Based	0.05638 (3.0194)	0.05638 (3.0194)	12.15%	2004-2016	443.84
Modified Approaches					
Momentum	0.01041 (0.5630)	0.01041 (0.5630)	10.31%	2004-2016	358.01
Back Testing	0.04236 (2.1591)	0.04236 (2.1591)	12.87%	2004-2016	482.35
Selectivity Alpha-Based	0.04141 (1.8968)	0.04617 (2.577)	12.82%	2004-2016	479.82
Selectivity Info-Ratio-Based	0.06251 (3.0167)	0.06251 (3.0167)	14.40%	2004-2016	574.62

Table 7 - Comparison of key results for all predictors

"Top Portfolio" reports alphas of the theoretically best portfolio of each predictor. "Largest Alpha" exhibits the largest alpha among all portfolios for each predictor. Corresponding t-statistics are provided in parentheses below each alpha. "Top Portfolio's CAGR" contains the compounded annual growth rate of the top portfolio of each predictor, while the column evaluation period depicts the time-period during which the respective portfolios record returns. The rightmost column depicts the payoff of an investment of SEK100 in the top portfolio, holding it for the entire evaluation period. All data is reported for both the original and the modified methodologies of each predictor.

[a] As stated in the column "Evaluation Period", the investment horizon for this approach spans over 2008 through 2016.

6.2. Comparison to Original Studies

The results we obtain provide partial support for the findings of the original papers, as two out of three predictors (back testing and selectivity predictors) prove to show predictive power in the Swedish mutual fund industry. When comparing our results to those of the original papers, one must keep in mind that they refer to different industries and time periods.

The results of the momentum predictor are disappointing to some extent. The alphas we obtain from both the original and modified approach are lower and less significant than in Hendricks et al. (1993).In their study, the 8^{th} octile portfolio always reports a large positive alpha, while the other portfolios report mainly negative alphas. However, Hendricks et al. (1993) run a CAPM-like regression for each portfolio on different benchmarks. The higher alphas, compared to our results, can be explained by the different regression model, as CAPM is shown to overestimate alphas by Fama and French (1993) due to the lack of common risk factors. All in all, we do not find that the original methodology of the momentum predictor leads to outperformance compared to Swedish market indexes. Our results do not show as clear trends throughout the portfolios as the original study, i.e. alphas are not monotonically increasing in octile. Furthermore, our modification seems to make the portfolios more vulnerable to interferences by autocorrelation and heteroscedasticity, as the modified approach leads to three portfolios instead of one that show signs of such interferences.

The application of the back testing predictor yields alphas of similar magnitude to the original study. Modifying the approach by reducing the estimation period to one year and using daily instead of monthly returns considerably improves the results, while also producing more significant alphas. Nevertheless, both the original and the modified approach cannot replicate the same performance patterns among the decile portfolios as the original study. In particular, alphas clearly decline linearly with increasing decile number in the original study, which is not the case in our study.[14] Strikingly, the Breusch-Pagan test shows clear signs of heteroscedasticity throughout the rank-portfolios from the modified approach. The extent of the difference between the original and the modified approach suggests that the change is associated with the use of daily data in the modified approach.

The selectivity predictor based on Amihud and Goyenko (2013) provides us with better results than those reported in the original study. We analyse the R^2s from an annual regression of daily fund excess returns on the Carhart four-factor returns for every fund and obtain 886 R^2s from 105 funds, ranging from 0.076 to 0.895. The mean R^2 is 0.575 and the median is 0.590. Amihud and Goyenko (2013) report higher average values for R^2, which may imply the regression model works better in their setting than ours. Even in our modified approach, which is based on more data, we report the same values. The difference between the studies is particularly pronounced when looking at the Info-Ratio sorted portfolios. In the original paper, the Info-Ratio-sorted portfolios report considerably lower alphas than the alpha-sorted ones, which is the opposite of what we find. Additionally, the difference in alphas in our approach is not as pronounced as in the original paper. The original study reports much more significant results on average. It is remarkable that most of the portfolios in our study exhibit positive alphas, while in the original study more than two thirds of the portfolios produce negative alphas. As this phenomenon is present in both the original and modified approaches, it is likely that the reason lies within our sample data.

[14] For reasons mentioned above, we employ a reversed portfolio ranking. In the original study, alpha declines with decreasing decile number, which is equivalent to the decreasing alphas with increasing decile number in our study.

6.3. General Considerations

Some considerations are common to all predictors. Although we are confident in our data sources, we point that we ran into some issues while preparing the databases for our studies. In particular, we note that some funds with international holdings naturally reported returns on days when Swedish markets were closed. This means we have to match the different databases that we used, namely fund returns from Bloomberg, benchmark returns from Thomson Reuters, USDSEK exchange rates and Swedish risk-free rates from the Riksbank and, finally, the Carhart four-factors from AQR. The matching is done using the Swedish trading calendar as a reference. Moreover, we disregard funds' reported returns in non-trading days and interpolate missing daily values for the USDSEK exchange rate and the risk-free rate.

Daily return data is usually noisier than monthly or quarterly data, because it deviates more from normality than less frequent observation data. In addition, Scholes and Williams (1977) find an econometric problem – they call it "nonsynchronous trading" – that is especially severe when using daily return data. It arises from the fact that reported closing prices usually do not represent actual closing prices, which in turn leads to erroneous variables and coefficients within regression models. We have observed distorted betas using daily data for the back testing and the selectivity predictor. Because falsely specified betas affect the estimates of alpha, both predictors are affected by the characteristics of daily return data. Nevertheless, the effect was more severe for the selectivity predictor, where alphas are actually used for the ranking process, which probably lowers the precision of the ranking process. While we are aware of the problems daily data might cause, we hold on to it because the construction of the specific predictors requires it. Furthermore, the outstanding performances of the best portfolios suggest that the effects of those problems are kept within reasonable limits, even though one should be cautious when interpreting the resulting alphas.

As the CRSP Survivorship Bias Free Mutual Fund Database for the US mutual fund industry, the dataset we use for Swedish mutual fund includes funds that stop reporting returns in some instances. This may be due to end of operations, merging or simple lack of reporting. Naturally, we know about these instances ex-post, which might lead to hindsight and survivorship bias. Thus, we decide to assign such funds an arbitrary return of -20%. We acknowledge that assigning an arbitrary return is not ideal. However, this is the most prudent approach out of the alternatives we discussed. This return is assumed for one period (be that one month, one quarter or one year), after which we assume the hypothetical investor would be aware of the disappearance of the fund and would divest. Through robustness tests, we verified that changing the assumed return of disappearing funds has a neglectable effect on the final portfolio's alpha.

Moreover, our conclusions based on CAGR values naturally depend on the benchmark indexes chosen for this purpose. Generally, we find that our results become more pronounced when comparing to the OMXS30 or the AQR Market Index. As these benchmarks perform worse than OMXSMCPI during the evaluation period, all the portfolios report higher alphas. Overall, the AQR Market Index seems to be the most comparable benchmark. This is especially true for the momentum predictor and the back testing predictor, at least in its original specification. As the performance of the other predictors is much higher, the OMXSMCPI is a closer benchmark for these three. These implications do not come as a surprise. The OMXS30 is limited to the 30 biggest companies in Sweden, with no loading on the size factor. The OMXSMCPI focuses on the Mid-Cap segment and clearly registers a premium compared to the main index. The AQR Market Index comes somewhat in the middle of the two, as it captures the performance of all listed Swedish companies.

One particular issue that might lead to the high alphas recorded from the selectivity and the modified back testing predictors is the model specification. The Carhart four-factor model might not be the most suitable to explain returns of mutual funds and, more specifically, of the presented predictors. Since we focus on equity funds, we assume that these risk factors are a reasonable choice. However, this specific question is an interesting opportunity for future research.

In general, the implementation of the Newey-West estimator increases the significance of our results. This is somewhat in contrast with the largely negative results from the Durbin-Watson and Breusch-Pagan tests. However, we note that the adjustments are larger in those cases where the tests are significant. One meaningful exception is represented by *Appendix 14* (alpha on TR^2), where the adjustments decrease the value of the t-statistics. Nonetheless, the values remain highly significant.

Finally, we point out that the results we derive are hypothetical, as they are only partly replicable by investors. While some portfolios do outperform their benchmarks, this performance may not be implementable by investors. First, our returns are gross of taxes, fees and front- and end-load. Additionally, most strategies assume a monthly holding period, which may be too short. This results in high turnover, which directly feeds into the fees we just referred to. Furthermore, some reported data might only be available with a certain time lag. Such a lag could potentially deteriorate the efficacy of the discussed predictor. Some funds also have minimum investment requirements, which may make it impossible for certain retail investors to invest in them. We leave these issues open for further research, as we were not able to retrieve data about fees and loads.

6.4. Predictor Specific Considerations

The results we obtain for the back testing predictor are puzzling in the sense that, when implementing our modified approach, all but one out of ten portfolios report positive alphas. Yet, even though those positive alphas are mostly not significant, we would expect more of the portfolios to report negative alphas, as the average mutual fund manager should not outperform the market. Looking at the striking performance of the modified predictor depicted in *Figure 4*, the modified methodology seems to work well in allocating the top performing funds to the 1^{st} decile portfolio.

Turning to the selectivity predictor, *Appendix 14* shows the summary of a regression of all alphas on the previous year's TR^2. While the negative coefficient on the independent variable is reassuring as it confirms the hypothesis of a negative relationship between R^2 and the following year's alpha or Info-Ratio, we note that the R^2s of the regression are lower than those reported by Amihud and Goyenko (2013). We believe this is due to two main factors. First, the number of observations, which is clearly higher in the original paper, probably has a substantial influence on R^2. We observe that the difference in R^2 between our study and the original paper is smaller when we use the modified approach that features more observations. Second, the fewer control variables most likely lead to a lower R^2. Due to the lack of fund specific data on the Swedish market, such as *Expenses, Manager Tenure, Fund Age or Turnover*, we do not include any control variables while analysing the selectivity predictor. As Amihud and Goyenko (2013) include several control variables in their regression analysis to account for their influence on mutual fund performance, our results are likely to be less precise in this regard. Including control variables would be useful in the analysis of all three performance predictors. However, it might also be the case that for our specific case other, nonlinear, model types might be a better fit, although testing for this would go beyond the scope of this study.

7. Conclusion

Over the past two decades, numerous academic papers have discovered statistical techniques to predict mutual fund performance, using mostly fund characteristics that are available to retail investors. Jones and Mo (2016) review most of these predictors and test whether they retain their predictive power out-of-sample. Even though predictive power decreases out-of-sample, Jones and Mo (2016) still report weak forms of it. Their results inspired us to provide another out-of-sample test of a smaller assortment of predictors in the Swedish mutual fund industry. Specifically, we focus on the possibility of using the predictors to achieve meaningful economic gains. Thus, the top performing portfolios for each predictor are our priority throughout the study. We include the following three predictors, which are all, theoretically, suitable for retail investors: a *momentum predictor* based on Hendricks et al. (1993), a *back testing predictor* based on Mamaysky et al. (2007) and a *selectivity predictor* based on Amihud and Goyenko (2013). We note that we both follow the methodologies used in the original papers as well as introduce our own modifications to each methodology, which mainly consist of different holding periods or estimation windows. We test the profitability of each approach using two different tools: first, we run Carhart (1997) four-factor regressions to test whether the methodologies can produce portfolios that exhibit positive alphas, controlling for common risk factors. Second, we plot the cumulative performance of the top portfolio, i.e. the best performing portfolio of funds by construction, against three Swedish equity benchmark indexes, namely the OMXS30, OMXSMCPI, and the AQR Market Index. We focus on actively managed Swedish equity mutual funds. Our sample period spans from January 2003 through December 2016.

Over the course of our study, we find that two out of three predictors show considerable predictive power. There are evident performance discrepancies between the different approaches. The selectivity predictor exhibits the best performance, both following the original and our modified methodology. The top portfolio clearly outperforms all three benchmark indexes throughout our sample period. The same observation holds for the top portfolio built from the modified back testing predictor. The original methodology of the back testing predictor and both approaches of the momentum predictor render merely mediocre to poor performances compared to the benchmarks. The results also reveal a clear pattern that the modified approaches outperform the original methodologies, both on the basis of the compounded annual growth rates of respective top portfolios as well as alphas from four-factor regressions. Moreover, throughout almost all modified approaches, top portfolios show positive and mostly highly significant alphas from regressing portfolio excess returns on the Carhart four-factors. The original approaches yield less significant but positive alphas. The momentum predictor behaves differently, with both top portfolios showing non-significant alphas, and a negative value in the original approach.

Our results partially provide out-of-sample confirmation of the findings of the original studies. Both the back testing methodology by Mamaysky et al. (2007) and the selectivity predictor by Amihud & Goyenko (2013) show out-of-sample predictive power. The alphas we obtain from the momentum predictor, both original and modified approach, are lower and less significant than those reported by Hendricks et al. (1993). Hence, we cannot confirm their results. While Hendricks et al. (2013) use the CAPM model, we implement Carhart's four-factor model to account for common risk factors, naturally lowering alphas.

It is important to note that all our data is gross of all fees and expenses. Hence, our results are inflated compared to what an investor could actually realise. However, the alphas we obtain, especially from the selectivity predictor and the modified back testing predictor would most likely still be substantial after accounting for such costs. Further, we have no access to a survivorship-free database for the Swedish market. Nevertheless, we apply measures during our data preparation and fund selection processes to mitigate this problem. We also adjust our results for partially present heteroscedasticity and autocorrelation interferences, using the Newey and West (1987) estimator.

Our study yields important findings for the following reasons: first, we provide an out-of-sample test of three established mutual fund performance predictors, partially confirming earlier findings. Second, the observation of well-functioning mutual fund predictors, theoretically deployable by any retail investor, is a valuable finding for investors. This may be of particular importance in the Swedish market, which is unique regarding its link to the Swedish pension system. The described predictors may provide investors and public pension scheme members alike with a decision guidance on which mutual funds to invest in. Third, our results partially promote arguments for investing in actively manged mutual funds, which are usually more expensive than index tracking passive funds. We show that certain strategies provide investors with returns that beat index tracking strategies.

References

Amihud, Y., and Goyenko, R., 2013, Mutual Fund's R2 as Predictor of Performance, *Review of Financial Studies*, 26(3), pp. 667–694.

AQR Capital Management, LLC, 2017,[online], *The Devil in HML's Details: Factors, Daily*, http://www.aqr.com/library/data-sets/the-devil-in-hmls-details-factors-daily/data, date of last access: 12/05/2017.

Asness, C., and Frazzini, A., 2013, The Devil in HML's Details, *The Journal of Portfolio Management*, 39(4), pp. 49–68.

Barras, L., Scaillet, O., and Wermers, R., 2010, False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas, *The Journal of Finance*, 65(1), pp. 179–216.

Brands, S., Brown, S.J., and Gallagher, D.R., 2005, Portfolio Concentration and Investment Manager Performance, *International Review of Finance*, 5(3-4), pp. 149–174.

Breusch, T.S., and Pagan, A.R., 1979, A Simple Test for Heteroscedasticity and Random Coefficient Variation, *Econometrica*, pp. 1287–1294.

Cao, C., Simin, T.T., and Wang, Y., 2013, Do Mutual Fund Managers Time Market Liquidity?, *Journal of Financial Markets*, 16(2), pp. 279–307.

Carhart, M.M., 1997, On Persistence in Mutual Fund Performance, *The Journal of Finance*, 52(1), pp. 57–82.

Chen, J., Hong, H., Huang, M., and Kubik, J.D., 2004, Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization, *The American Economic Review*, 94(5), pp. 1276–1302.

Christoffersen, S.E., and Sarkissian, S., 2009, City Size and Fund Performance, *Journal of Financial Economics*, 92(2), pp. 252–275.

Cohen, R.B., Coval, J.D., and Pástor, L., 2005, Judging Fund Managers by the Company They Keep, *The Journal of Finance*, 60(3), pp. 1057–1096.

Cremers, K.M., and Petajisto, A., 2009, How Active is Your Fund Manager? A new Measure that Predicts Performance, *Review of Financial Studies*, 22(9), pp. 3329–3365.

Da, Z., Gao, P., and Jagannathan, R., 2011, Impatient Trading, Liquidity Provision, and Stock Selection by Mutual Funds, *Review of Financial Studies*, 24(3), pp. 675–720.

Daniel, K., Grinblatt, M., Titman, S., and Wermers, R., 1997, Measuring Mutual Fund Performance with Characteristic-Based Benchmarks, *The Journal of Finance*, 52(3), pp. 1035–1058.

Durbin, J., and Watson, G.S., 1951, Testing for Serial Correlation in Least Squares Regression. II, *Biometrika*, 38(1–2), pp. 159–178.

Easley, D., Kiefer, N.M., O'hara, M., and Paperman, J.B. 1996, Liquidity, Information, and Infrequently Traded Stocks, *The Journal of Finance*, 51(4), pp. 1405–1436.

Fama, E.F., and French, K.R., 1993, Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics*, 33(1), pp. 3–56.

Goetzmann, W.N., and Ibbotson, R.G., 1994, Do Winners Repeat?, *The Journal of Portfolio Management*, 20(2), pp. 9–18.

Grinblatt, M., and Titman, S., 1989a, Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings, *The Journal of Business*, pp. 393–416.

Grinblatt, M., and Titman, S., 1989b, Portfolio Performance Evaluation: Old Issues and New Insights, *Review of Financial Studies*, 2(3), pp. 393–421.

Grinblatt, M., and Titman, S., 1992, The Persistence of Mutual Fund Performance, *The Journal of Finance*, 47(5), pp. 1977–1984.

Hendricks, D., Patel, J., and Zeckhauser, R., 1993, Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance, 1974–1988, *The Journal of Finance*, 48(1), pp. 93–130.

Huang, J., Sialm, C., and Zhang, H., 2011, Risk Shifting and Mutual Fund Performance, *Review of Financial Studies*, 24(8), pp. 2575–2616.

Jensen, M.C., 1968, The Performance of Mutual Funds in the Period 1945–1964, *The Journal of Finance*, 23(2), pp. 389–416.

Jones, C.S., and Mo, H., 2016, Out-of-Sample Performance of Mutual Fund Predictors, *working paper, University of Southern California*.

Kacperczyk, M., van Nieuwerburgh, S., and Veldkamp, L., 2014, Time-Varying Fund Manager Skill, *The Journal of Finance*, 69(4), pp. 1455–1484.

Kacperczyk, M., and Seru, A., 2007, Fund Manager Use of Public Information: New Evidence on Managerial Skills, *The Journal of Finance*, 62(2), pp. 485–528.

Kacperczyk, M., Sialm, C., and Zheng, L., 2005, On the Industry Concentration of Actively Managed Equity Mutual Funds, *The Journal of Finance*, 60(4), pp. 1983–2011.

Kacperczyk, M., Sialm, C., and Zheng, L., 2008, Unobserved Actions of Mutual Funds, *Review of Financial Studies*, 21(6), pp. 2379–2416.

Kosowski, R., Timmermann, A., Wermers, R., and White, H., 2006, Can Mutual Fund "Stars" Really Pick Stocks? New Evidence from a Bootstrap Analysis, *The Journal of Finance*, 61(6), pp. 2551–2595.

Malkiel, B.G., 1995, Returns from Investing in Equity Mutual Funds 1971 to 1991, *The Journal of Finance*, 5(2), pp. 549–572.

Mamaysky, H., Spiegel, M., and Zhang, H., 2007, Improved Forecasting of Mutual Fund Alphas and Betas, *Review of Finance*, 11(3), pp. 359–400.

Newey, W.K., and West, K.D., 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelationconsistent Covariance Matrix, *Econometrica*, 55(3), pp. 703–708.

Scholes, M., and Williams, J., 1977, Estimating Betas from Nonsynchronous Data, *Journal of Financial Economics*, 5(3), pp. 309–327.

Sharpe, W.F., 1966, Mutual Fund Performance, *The Journal of Business*, 39(1), pp. 119–138.

Treynor, J., and Mazuy, K., 1966, Can Mutual Funds Outguess the Market?, *Harvard Business Review*, 44(4), pp. 131–136.

Treynor, J.L., and Black, F., 1973, How to Use Security Analysis to Improve Portfolio Selection, *The Journal of Business*, 46(1), pp. 66–86.

White, H., 1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica*, 48(4), pp. 817–838.

APPENDIX

Summary Statistics of the Sorting Variable for the Momentum Predictor (Original Approach)

Portfolio Octile	Minimum	Maximum	Mean	Standard Deviation
1	-25.15%	26.22%	0.89%	9.26%
2	-23.27%	29.35%	1.89%	9.16%
3	-23.29%	28.00%	2.14%	8.87%
4	-22.82%	27.65%	2.48%	9.13%
5	-20.81%	28.17%	2.83%	9.00%
6	-19.01%	27.38%	3.24%	8.85%
7	-22.30%	30.62%	3.95%	9.68%
8	-18.71%	30.99%	5.47%	9.53%

Appendix 1 - Summary statistics of the sorting variable for the momentum predictor (original approach)
The sorting variable for the momentum predictor is the fund's performance in the year before the construction of the portfolios. Funds are ranked by performance and assigned to octiles, where the best performing funds are assigned to octile 8. We take the average of the funds' lagged performances in each period for each octile and report summary statistics for it. Minimum represents the lowest average performance recorded by funds before entering a portfolio, maximum presents the highest value of the average performance. We report the mean and standard deviation of these average performances. Figures in the table are, accordingly to the approach, expressed quarterly.

Summary Statistics of the Sorting Variable for the Momentum Predictor (Modified Approach)

Portfolio Octile	Minimum	Maximum	Mean	Standard Deviation
1	-20.58%	24.79%	0.28%	4.93%
2	-19.89%	23.81%	0.59%	4.98%
3	-18.24%	22.60%	0.71%	4.90%
4	-18.60%	23.91%	0.81%	4.94%
5	-17.87%	24.15%	0.92%	5.00%
6	-17.70%	21.95%	1.18%	4.81%
7	-16.15%	25.29%	1.29%	5.00%
8	-14.87%	24.71%	1.71%	5.01%

Appendix 2 - Summary statistics of the sorting variable for the momentum predictor (modified approach)
The sorting variable for the momentum predictor is the fund's performance in the year before the construction of the portfolios. Funds are ranked by performance and assigned to octiles, where the best performing funds are assigned to octile 8. We take the average of the funds' lagged performances in each period for each octile and report summary statistics for it. Minimum represents the lowest average performance recorded by funds before entering a portfolio, maximum presents the highest value of the average performance. We report the mean and standard deviation of these average performances. Figures in the table are, accordingly to the approach, expressed monthly.

Summary Statistics of the Sorting Variable for the Back testing Predictor (original approach)

Portfolio Decile	Minimum	Maximum	Mean	Standard Deviation
1	-0.00042	0.01035	0.00505	0.00210
2	-0.00094	0.00596	0.00324	0.00189
3	-0.00118	0.00530	0.00230	0.00183
4	-0.00149	0.00498	0.00161	0.00178
5	-0.00175	0.00477	0.00105	0.00163
6	-0.00225	0.00451	0.00048	0.00148
7	-0.00259	0.00354	-0.00005	0.00140
8	-0.00386	0.00156	-0.00054	0.00138
9	-0.00401	0.00117	-0.00115	0.00142
10	-0.00600	0.00083	-0.00208	0.00161

Appendix 3 - Summary statistics of the sorting variable for the back testing predictor (original approach)
The sorting variable for the back testing predictor is the estimated alpha of funds that are in the active pool after two filters are applied. The alphas are obtained from a regression of each fund's monthly excess returns on the Carhart four-factors over the previous five years. The table reports summary statistics for the average alpha recorded by the second-stage regression. Reported alphas are, accordingly to the approach, monthly. We record nine instances of no portfolios successfully passing both steps of the back testing procedure.

Summary Statistics of the Sorting Variable for the Back testing Predictor (modified approach)

Portfolio Decile	Minimum	Maximum	Mean	Standard Deviation
1	-0.00057	0.00210	0.00073	0.00053
2	-0.00072	0.00160	0.00045	0.00049
3	-0.00078	0.00147	0.00039	0.00047
4	-0.00084	0.00137	0.00032	0.00046
5	-0.00091	0.00131	0.00027	0.00044
6	-0.00096	0.00121	0.00023	0.00045
7	-0.00099	0.00117	0.00018	0.00044
8	-0.00106	0.00107	0.00012	0.00044
9	-0.00123	0.00101	0.00006	0.00044
10	-0.00141	0.00080	-0.00006	0.00047

Appendix 4 - Summary statistics of the sorting variable for the back testing predictor (modified approach)
The sorting variable for the back testing predictor is the estimated alpha of funds that are in the active pool after two filters are applied. The alphas are obtained from a regression of each fund's daily excess returns on the Carhart four-factors over the previous year. The table reports summary statistics for the average alpha recorded by the second-stage regression. Reported alphas are, accordingly to the approach, daily. We record 112 instances of no portfolios successfully passing both steps of the back testing procedure.

Summary Statistics of the Sorting Variable for the selectivity Predictor (original approach, alpha-sorted)

TR^2 Quartile	Alpha Quartile	Minimum	Maximum	Mean	Standard Deviation
1	1	0.53231	0.79782	0.69780	0.08503
		{-0.1737}	{0.1660}	{0.0126}	{0.1004}
	2	0.48131	0.79898	0.69931	0.09433
		{-0.1415}	{0.1857}	{0.0451}	{0.1025}
	3	0.49762	0.81986	0.69632	0.09358
		{-0.1153}	{0.2048}	{0.0679}	{0.1053}
	4	0.52586	0.80630	0.68887	0.09004
		{-0.0817}	{0.2662}	{0.1087}	{0.1062}
2	1	0.41707	0.73563	0.60781	0.09428
		{-0.2068}	{0.1719}	{0.0097}	{0.1119}
	2	0.41533	0.74744	0.61101	0.09474
		{-0.1747}	{0.2110}	{0.0383}	{0.1075}
	3	0.38906	0.73874	0.60847	0.10077
		{-0.1488}	{0.2281}	{0.0567}	{0.1039}
	4	0.42851	0.74087	0.60730	0.09422
		{-0.1171}	{0.2772}	{0.0879}	{0.1072}
3	1	0.34674	0.68918	0.55070	0.09103
		{-0.2475}	{0.1824}	{0.0008}	{0.1269}
	2	0.35918	0.68430	0.54964	0.08755
		{-0.1804}	{0.2135}	{0.0495}	{0.1216}
	3	0.35847	0.67891	0.55516	0.09078
		{-0.1080}	{0.2439}	{0.0920}	{0.1172}
	4	0.34028	0.67390	0.53445	0.10018
		{-0.0604}	{0.3351}	{0.1468}	{0.1216}
4	1	0.24762	0.59384	0.43443	0.09360
		{-0.2304}	{0.2118}	{0.0294}	{0.1399}
	2	0.21744	0.58783	0.43928	0.10434
		{-0.1757}	{0.2808}	{0.0932}	{0.1411}
	3	0.23311	0.58506	0.42336	0.10338
		{-0.1090}	{0.3065}	{0.1371}	{0.1385}
	4	0.18550	0.51026	0.38851	0.08995
		{-0.0012}	{0.3873}	{0.2122}	{0.1290}

Appendix 5 - Summary statistics of the sorting variable for the selectivity predictor, original approach with sorting on alpha

The sorting variables for the selectivity predictor are the TR^2, where $TR^2 = \log(\sqrt{(R^2/(1-\sqrt{(R^2)}))})$, and the alpha from a regression of daily fund's excess returns on the Carhart-four-factors over a year. The procedure is repeated every year. Funds are ranked by descending R^2 and then by ascending alpha. The table reports the average TR^2, and alpha in brackets, of the funds that compose each portfolio. The frequency in this approach is yearly.

Summary Statistics of the Sorting Variable for the selectivity Predictor (original approach, Info-Ratio-sorted)

TR^2 Quartile	Info-Ratio Quartile	Minimum	Maximum	Mean	Standard Deviation
1	1	0.53231	0.80897	0.70098	0.09259
		{-0.0948}	{0.1527}	{0.0166}	{0.0671}
	2	0.48131	0.79804	0.69505	0.09188
		{-0.0697}	{0.1693}	{0.0381}	{0.0686}
	3	0.46083	0.80286	0.68964	0.09757
		{-0.0636}	{0.1944}	{0.0531}	{0.0734}
	4	0.53304	0.82309	0.69703	0.08366
		{-0.0469}	{0.2137}	{0.0805}	{0.0777}
2	1	0.41487	0.73943	0.60933	0.09618
		{-0.0841}	{0.1252}	{0.0144}	{0.0594}
	2	0.39972	0.74364	0.60832	0.10060
		{-0.0726}	{0.1375}	{0.0303}	{0.0591}
	3	0.40687	0.74048	0.60970	0.09562
		{-0.0622}	{0.1517}	{0.0409}	{0.0603}
	4	0.42851	0.73913	0.60723	0.09171
		{-0.0514}	{0.1910}	{0.0609}	{0.0671}
3	1	0.36528	0.68918	0.54952	0.08980
		{-0.0969}	{0.1501}	{0.0105}	{0.0677}
	2	0.35398	0.67889	0.55161	0.08837
		{-0.0746}	{0.1740}	{0.0378}	{0.0664}
	3	0.35769	0.67715	0.55335	0.08915
		{-0.0425}	{0.1885}	{0.0607}	{0.0692}
	4	0.32773	0.68108	0.53576	0.10194
		{-0.0245}	{0.2477}	{0.0917}	{0.0798}
4	1	0.19249	0.59384	0.42233	0.10884
		{-0.0957}	{0.1475}	{0.0227}	{0.0678}
	2	0.22127	0.58783	0.43889	0.10218
		{-0.0537}	{0.2018}	{0.0620}	{0.0774}
	3	0.21959	0.56143	0.42799	0.09037
		{-0.0398}	{0.2358}	{0.0908}	{0.0853}
	4	0.20352	0.52127	0.39953	0.08647
		{-0.0041}	{0.2769}	{0.1327}	{0.0930}

Appendix 6 - Summary statistics of the sorting variable for the selectivity predictor, original approach with sorting on Info-Ratio

The sorting variables for the selectivity predictor are the TR^2, where $TR^2 = \log(\sqrt{(R^2/(1-\sqrt{(R^2)}))})$, and the Info-Ratio from a regression of daily fund's excess returns on the Carhart-four-factors over a year. The procedure is repeated every year. Funds are ranked by descending R^2 and then by ascending Info-Ratio. The table reports the average TR^2, and alpha in brackets, of the funds that compose each portfolio. The frequency in this approach is yearly.

Summary Statistics of the Sorting Variable for the selectivity Predictor (modified approach, alpha-sorted)

TR^2 Quartile	Alpha Quartile	Minimum	Maximum	Mean	Standard Deviation
1	1	0.46058	0.83507	0.69624	0.09607
		{-0.2470}	{0.1839}	{0.0115}	{0.0935}
	2	0.46198	0.83678	0.68518	0.09848
		{-0.2263}	{0.2462}	{0.0452}	{0.0933}
	3	0.46823	0.84025	0.68384	0.09698
		{-0.2030}	{0.2722}	{0.0674}	{0.0941}
	4	0.45942	0.85880	0.68521	0.10141
		{-0.1721}	{0.3225}	{0.1030}	{0.0958}
2	1	0.39509	0.76936	0.59968	0.09567
		{-0.2926}	{0.2271}	{0.0101}	{0.1016}
	2	0.38035	0.77975	0.60004	0.09808
		{-0.2546}	{0.2728}	{0.0397}	{0.0966}
	3	0.38808	0.77765	0.59990	0.09841
		{-0.2296}	{0.2953}	{0.0591}	{0.0958}
	4	0.40543	0.77794	0.60099	0.09651
		{-0.1836}	{0.3384}	{0.0939}	{0.0976}
3	1	0.34700	0.74896	0.54731	0.09340
		{-0.3217}	{0.2553}	{0.0134}	{0.1096}
	2	0.34343	0.72874	0.54525	0.09108
		{-0.2523}	{0.3048}	{0.0491}	{0.1075}
	3	0.33569	0.72470	0.53639	0.09636
		{-0.2140}	{0.3480}	{0.0810}	{0.1091}
	4	0.32643	0.74710	0.53027	0.10012
		{-0.1361}	{0.4466}	{0.1299}	{0.1109}
4	1	0.16061	0.62676	0.42155	0.10413
		{-0.3162}	{0.3130}	{0.0298}	{0.1309}
	2	0.18153	0.64604	0.43381	0.10346
		{-0.2426}	{0.3943}	{0.0947}	{0.1316}
	3	0.15925	0.60811	0.42646	0.10364
		{-0.1782}	{0.4200}	{0.1337}	{0.1291}
	4	0.14565	0.60911	0.38635	0.10610
		{-0.0787}	{0.4508}	{0.2054}	{0.1208}

Appendix 7 - Summary statistics of the sorting variable for the selectivity predictor, original approach with sorting on alpha

The sorting variables for the selectivity predictor are the TR^2, where $TR^2 = \log(\sqrt{(R^2/(1-\sqrt{(R^2)}))})$, and the alpha from a regression of daily fund's excess returns on the Carhart-four-factors over a year. Funds are ranked by descending R^2 and then by ascending Info-Ratio. The procedure is repeated every month. The table reports the average TR^2, and alpha in brackets, of the funds that compose each portfolio. The frequency in this approach is monthly.

Summary Statistics of the Sorting Variable for the selectivity Predictor (modified approach, Info-Ratio-sorted)

TR^2 Quartile	Alpha Quartile	Minimum	Maximum	Mean	Standard Deviation
1	1	0.46058	0.83507	0.69296	0.10011
		{-0.1398}	{0.1635}	{0.0144}	{0.0625}
	2	0.46198	0.83678	0.68314	0.09945
		{-0.1109}	{0.1697}	{0.0368}	{0.0616}
	3	0.46083	0.84025	0.68811	0.09539
		{-0.1057}	{0.1984}	{0.0519}	{0.0658}
	4	0.45942	0.85612	0.68625	0.09763
		{-0.0955}	{0.2335}	{0.0752}	{0.0706}
2	1	0.39185	0.76936	0.59962	0.09716
		{-0.1307}	{0.1402}	{0.0146}	{0.0559}
	2	0.38601	0.77975	0.59978	0.09818
		{-0.1155}	{0.1553}	{0.0308}	{0.0560}
	3	0.39206	0.77765	0.59985	0.09672
		{-0.1054}	{0.1967}	{0.0426}	{0.0589}
	4	0.39696	0.77794	0.60144	0.09667
		{-0.0859}	{0.2618}	{0.0647}	{0.0648}
3	1	0.34561	0.74896	0.54560	0.09435
		{-0.1455}	{0.1501}	{0.0149}	{0.0582}
	2	0.33826	0.73559	0.54513	0.09275
		{-0.1128}	{0.1724}	{0.0350}	{0.0601}
	3	0.34309	0.72470	0.53794	0.09286
		{-0.0977}	{0.2388}	{0.0538}	{0.0678}
	4	0.32173	0.74710	0.53048	0.10185
		{-0.0604}	{0.3105}	{0.0850}	{0.0779}
4	1	0.16061	0.62676	0.42279	0.10682
		{-0.1517}	{0.1682}	{0.0242}	{0.0659}
	2	0.17457	0.64604	0.42447	0.10934
		{-0.1038}	{0.2520}	{0.0619}	{0.0744}
	3	0.20784	0.61288	0.42899	0.09727
		{-0.0705}	{0.2850}	{0.0867}	{0.0811}
	4	0.11173	0.61427	0.39303	0.10986
		{-0.0213}	{0.3276}	{0.1249}	{0.0892}

Appendix 8 - Summary statistics of the sorting variable for the selectivity predictor, original approach with sorting on Info-Ratio

The sorting variables for the selectivity predictor are the TR^2, where $TR^2 = \log(\sqrt{(R^2/(1-\sqrt{(R^2)}))})$, and the Info-Ratio from a regression of daily fund's excess returns on the Carhart-four-factors over a year. The procedure is repeated every month. Funds are ranked by descending R^2 and then by ascending Info-Ratio. The table reports the average TR^2, and alpha in brackets, of the funds that compose each portfolio. The frequency in this approach is monthly.

Original Momentum Predictor (Unadjusted)

	Alpha	AQR Market Index	SMB	HML	UMD
Octile 1	-0.00229	0.9066	0.0086	0.1435	-0.1350
	(-0.1647)				
Octile 2	-0.02234	0.9384	0.0189	0.0220	-0.0330
	(-1.3821)				
Octile 3	-0.00507	0.9872	0.0761	0.0028	-0.0827
	(-0.3433)				
Octile 4	0.00148	0.9844	-0.0093	0.0329	-0.0621
	(0.1145)				
Octile 5	-0.01369	1.0047	0.0016	0.1106	-0.0993
	(-1.0469)				
Octile 6	-0.02746	1.0570	0.0590	-0.1240	0.0503
	(-1.6891)				
Octile 7	-0.00362	1.0530	0.0949	-0.0783	-0.0204
	(-0.1989)				
Octile 8	-0.01676	1.0596	0.0974	-0.0944	0.0662
	(-0.7668)				

Appendix 9 - Summary statistics from a Carhart four-factor regression of portfolios' excess returns– original approach of the momentum predictor

In every period, funds are sorted into octiles based on their performance during the previous four quarters and portfolios are built based on the octiles (a higher octile corresponds to a higher past performance). Each portfolio is held for one quarter, then the procedure is repeated. The first portfolio is built in 1Q2004 and the last in 3Q2016. We run a regression of each portfolio's excess returns on the four Carhart factors and report alphas and betas to each factor. The regression is based on quarterly data but alphas are annualised. T-statistics are reported in brackets below each alpha. P-values from statistical tests, checking for autocorrelation and heteroscedasticity, are also reported.

Modified Momentum Predictor (Unadjusted)

	Alpha	AQR Market Index	SMB	HML	UMD
Octile 1	0.00579	0.8809	0.1036	0.0184	-0.1047
	(0.3737)				
Octile 2	-0.01626	0.9185	0.0492	-0.0785	-0.0277
	(-1.0561)				
Octile 3	-0.00779	0.9561	0.1213	-0.0519	-0.1157
	(-0.5498)				
Octile 4	0.00821	0.9505	0.0681	-0.0235	-0.0793
	(0.6000)				
Octile 5	0.00289	0.9445	0.0909	-0.0140	-0.0930
	(0.1916)				
Octile 6	0.00251	0.9234	0.1037	-0.0305	-0.0722
	(0.1570)				
Octile 7	0.00355	0.9536	0.1603	-0.0948	-0.0635
	(0.2138)				
Octile 8	0.01041	0.9154	0.1372	-0.1236	-0.0144
	(0.5574)				

Appendix 10 - Summary statistics from a Carhart four-factor regression of portfolios' excess returns– modified approach of the momentum predictor

In every period, funds are sorted into octiles based on their performance during the previous twelve months and portfolios are built based on the octiles (a higher octile corresponds to a higher past performance). Each portfolio is held for one month, then the procedure is repeated. The first portfolio is built in January 2004 and the last in November 2016. We run a regression of each portfolio's excess returns on the four Carhart factors and report alphas and betas to each factor. The regression is based on monthly data but alphas are annualised. T-statistics are reported in brackets below each alpha. P-values from statistical tests, checking for autocorrelation and heteroscedasticity, are also reported.

Original Back Testing Predictor (Unadjusted)

	Alpha	AQR Market Index	HML	SMB	UMD
Decile 1	0.02008	0.9810	0.2025	-0.1450	-0.0989
	(0.7953)				
Decile 2	0.02927	0.8880	0.1974	-0.1379	-0.1142
	(0.9238)				
Decile 3	0.01212	0.8166	0.0306	-0.0790	0.0012
	(0.4725)				
Decile 4	0.03276	0.9351	0.0889	-0.1127	-0.0530
	(1.4613)				
Decile 5	0.02335	0.9380	0.0813	-0.0985	-0.0947
	(1.0700)				
Decile 6	0.00002	0.9675	0.1441	-0.0501	-0.1265
	(0.0007)				
Decile 7	-0.00278	0.8854	0.1150	0.0206	-0.1188
	(-0.1045)				
Decile 8	-0.00682	0.9130	0.0598	-0.1050	-0.0557
	(-0.3164)				
Decile 9	-0.03441	0.9805	0.0645	-0.1123	-0.0143
	(-1.2755)				
Decile 10	-0.01940	0.9601	0.2251	-0.0510	-0.1220
	(-0.6558)				

Appendix 11 - Summary statistics from a Carhart four-factor regression of portfolios' excess returns – original approach of the back testing predictor

Every month, funds are filtered through a two-step back test. The first step compares the estimated alpha to the next realised alpha. The estimated alpha is from a Carhart four-factor regression of each fund's returns, based on 60 monthsof data. The alpha from this regression is compared to the realised excess return in month 61 (fund's return in excess of market return) and, if the sign of the two is the same, the fund stays in the active pool. The next step involves a second Carhart four-factor regression, with a one-month forward shift relative to the first step. If estimates of alpha and beta (relative to the Market factor) fall within a specified interval, the fund remains in the active pool. Funds in the active pool are ranked by alpha into deciles. We employ a reversed ranking which results in the "best performing funds populating the 1st decile. Equally-weighted portfolios are built based on the deciles and are held for one month, after which the procedure is repeated. Due to the five-year estimation window, the first decile-based portfolios are built in January 2008, while the last ones are built in December 2016. Alphas and betas are from a Carhart four-factor regression of the portfolios' excess returns. Alphas are annualised as the regression yields monthly alphas. T-statistics are presented in parentheses below the according alpha. P-values from statistical tests, checking for autocorrelation and heteroscedasticity, are also reported.

Modified Back Testing Predictor (Unadjusted)

	Alpha	AQR Market Index	HML	SMB	UMD
Decile 1	0.04236	0.9739	0.1240	-0.0767	-0.0809
	(1.8982)				
Decile 2	0.01004	0.8803	0.0552	-0.0201	-0.0031
	(0.4735)				
Decile 3	0.02555	0.8649	0.1502	-0.0904	-0.0147
	(1.1357)				
Decile 4	0.00651	0.8288	0.1234	-0.0015	-0.1300
	(0.2841)				
Decile 5	0.01040	0.8375	0.1819	-0.0853	-0.0572
	(0.4288)				
Decile 6	0.00479	0.9001	0.1391	-0.0411	-0.0967
	(0.2675)				
Decile 7	0.01335	0.8644	0.1081	0.0217	-0.1017
	(0.5821)				
Decile 8	0.01199	0.8121	0.2338	-0.0485	-0.0777
	(0.5366)				
Decile 9	-0.00487	0.8717	0.1316	-0.0113	-0.1267
	(-0.2315)				
Decile 10	0.00193	0.7092	0.2678	0.0270	-0.1233
	(0.0673)				

Appendix 12 - Summary statistics from a Carhart four-factor regression of portfolios' excess returns – modified approach of the back testing predictor

Every month, funds are filtered through a two-step back test. The first step compares the estimated alpha to the next realised alpha. The estimated alpha is from a Carhart four-factor regression of each fund's excess returns, based on twelve months' of daily data. The alpha from this regression is compared to the realised excess return in month 61 (fund's return over market return) and, if the sign of the two is the same, the fund stays in the active pool. The next step involves a second Carhart four-factor regression, with a one-month forward shift relative to the first step. If estimates of alpha and beta (relative to the Market factor) fall within a specified interval, the fund remains in the active pool. Funds in the active pool are ranked by alpha into deciles. We employ a reversed ranking which results in the best performing funds populating the 1st decile. Equally-weighted portfolios are built based on the deciles and are held for one month, after which the procedure is repeated. Due to the one-year estimation window, the first decile-based portfolios are built in January 2004, while the last ones are built in December 2016. Alphas and betas are from a Carhart four-factor regression of the portfolios' excess returns. Alphas are annualised as the regression yields monthly alphas. T-statistics are presented in parentheses below the according alpha. P-values from statistical tests, checking for autocorrelation and heteroscedasticity, are also reported.

Selectivity Predictor (Unadjusted)

Original Approaches

R^2_{t-1}	Alpha$_{t-1}$ Low	2	3	High
Low	0.00051	0.01345	0.00982	-0.00171
	(0.0385)	(1.0198)	(0.7736)	(-0.1250)
2	0.00402	0.00111	-0.00416	-0.00163
	(0.2803)	(0.0752)	(-0.3085)	(-0.1138)
3	0.02232	-0.00025	-0.00615	0.02808
	(1.2578)	(-0.0154)	(-0.4001)	(1.6426)
High	0.03210	0.03413	0.04600	0.04406
	(1.7011)	(1.7273)	(2.2136)	(2.1813)

R^2_{t-1}	Info-Ratio$_{t-1}$ Low	2	3	High
Low	-0.00047	0.00713	0.01695	-0.00350
	(-0.0369)	(0.5374)	(1.1731)	(-0.2602)
2	0.00915	0.00015	-0.00789	-0.00141
	(0.6111)	(0.0109)	(-0.5609)	(-0.0987)
3	0.02016	-0.00655	0.00604	0.02278
	(1.1355)	(-0.3989)	(0.3942)	(1.3594)
High	0.02648	0.02962	0.04554	0.05638
	(1.4092)	(1.4745)	(2.2137)	(2.7612)

Modified Approaches

R^2_{t-1}	Alpha$_{t-1}$ Low	2	3	High
Low	0.00516	0.00768	0.00215	0.01285
	(0.3724)	(0.5999)	(0.1616)	(0.9200)
2	-0.00835	-0.00211	-0.00085	-0.00564
	(-0.5587)	(-0.1373)	(-0.0613)	(-0.4078)
3	0.00610	0.00780	0.00556	0.02545
	(0.3981)	(0.4680)	(0.3196)	(1.6135)
High	0.04011	0.04617	0.04279	0.04141
	(2.2973)	(2.3148)	(2.2453)	(1.8349)

R^2_{t-1}	Info-Ratio$_{t-1}$ Low	2	3	High
Low	0.01115	0.00279	0.00079	0.01471
	(0.7897)	(0.2163)	(0.0620)	(1.0831)
2	-0.00587	-0.00027	-0.00541	-0.00427
	(-0.3874)	(-0.0173)	(-0.3772)	(-0.3101)
3	0.00461	0.00470	0.00756	0.02835
	(0.2842)	(0.2935)	(0.4420)	(1.7293)
High	0.03754	0.03000	0.04480	0.06251
	(2.0181)	(1.5892)	(2.2421)	(2.8247)

Appendix 13 - Summary statistics from a Carhart four-factor regression of portfolio returns – selectivity predictor

In every period (every year in the original approach, every month in the modified approach), each fund's daily excess returns are regressed on the four Carhart factors. The regression spans over one year of historical data. The R^2 of the regression is retained and transformed into TR^2. $TR^2 = \log(\sqrt{R^2/(1-\sqrt{R^2})})$. Funds are ranked into quartiles by TR^2 and, within these quartiles, by alpha (Info-Ratio). TR^2 are reversely ordered, so that quartile 4 includes the funds with the lowest TR^2s, which should perform better. Each portfolio is held for one year (original approach) or one month (modified approach). When the holding period expires, the procedure is repeated. We compute excess returns for the portfolios and regress them on the four Carhart factors. We note that in the original approach we calculate monthly excess returns to use in the regression, to increase the number of data points. We report annualised alphas for each portfolio and, below them, Newey-West adjusted t-statistics.

Selectivity Predictor

Variables	Original Approaches		Modified Approaches	
	Alpha	Info-Ratio	Alpha	Info-Ratio
Newey-West Adjusted Output				
TR^2	-0.05722	-0.03727	-0.08508	-0.05614
T-statistics	-8.43650	-8.6127	-13.6810	-12.4670
P-Value	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16
Intercept	0.14119	0.09637	0.16515	0.11204
T-statistics	16.0867	16.6034	21.4240	19.0250
P-Value	<2.2e-16	<2.2e-16	<2.2e-16	<2.2e-16
R^2	0.0597	0.0697	0.1538	0.1793
Degrees of Freedom	850	850	9397	9397
Durbin-Watson P-Value	0.01449	0.00006	< 2.2e-16	< 2.2e-16
Breusch-Pagan P-Value	0.75310	0.00287	0.00009	< 2.2e-16
Output without Newey-West Adjustment				
TR^2	-0.05722	-0.03727	-0.08508	-0.05614
T-statistics	-7.34	-7.98	-41.32	-45.31
P-Value	4.84E-13	4.76E-15	< 2.2e-16	< 2.2e-16
Alpha	0.14119	0.09637	0.16515	0.11204
T-statistics	13.99	15.93	63.53	71.61
P-Value	< 2.2e-16	< 2.2e-16	< 2.2e-16	< 2.2e-16

Appendix 14 - Annual and monthly regression for the original and the modified approach of the selectivity predictor

Fund alphas or info-ratios of period t are regressed on the TR²s of the previous year. $TR^2 = \log(\sqrt{(R^2/(1-\sqrt{(R^2)}))})$, where R^2 is obtained from an annual regression of daily fund excess returns on the Carhart four-factors. Info-Ratio=alpha/RMSE, where RMSE is the root mean squared error from the annual regression. P-values of the Durbin-Watson and Breusch-Pagan Tests are reported to test for autocorrelation and heteroscedasticity. The upper part of the table reports Newey-West adjusted values, while the unadjusted values can be found at the bottom of table.

Durbin-Watson Test P-Values

Original Approach

R^2_{t-1}	Alpha$_{t-1}$ Low	2	3	High
Low	0.99944	0.91706	0.99925	0.99938
2	0.99206	0.99994	0.99999	0.99982
3	0.99781	1.00000	0.99994	0.96571
High	0.91031	0.79646	0.98115	0.92338

R^2_{t-1}	Info-Ratio$_{t-1}$ Low	2	3	High
Low	0.99848	0.97950	0.99950	0.89365
2	0.98285	0.99990	0.99979	0.99999
3	0.99564	0.99999	0.99997	0.95885
High	0.87889	0.88024	0.90184	0.97352

Modified Approach

R^2_{t-1}	Alpha$_{t-1}$ Low	2	3	High
Low	0.99977	0.99858	0.99857	0.94974
2	0.99997	0.99983	0.99999	0.99799
3	0.99808	1.00000	0.99964	0.98052
High	0.99503	0.67104	0.98226	0.43409

R^2_{t-1}	Info-Ratio$_{t-1}$ Low	2	3	High
Low	0.99993	0.98518	0.99975	0.95084
2	0.99986	0.99976	0.99999	0.99494
3	0.99642	0.99998	0.99973	0.97113
High	0.92812	0.74879	0.96508	0.70748

Appendix 15 - P-Values from Durbin-Watson test for all four methodologies for the selectivity predictor
Each portfolio we construct is tested for autocorrelation using the Durbin-Watson test. Values in the matrices are p-values of the test outcome. The test is based on a null hypothesis stating that no autocorrelation is present. Since all p-values reported are rather high and nowhere near significance, we cannot reject the null hypothesis. Hence, the test output suggests that, even though we cannot be certain, autocorrelation is not an issue on a portfolio level.

Breusch-Pagan Test P-Values

Original Approach

R^2_{t-1}	Alpha$_{t-1}$ Low	2	3	High
Low	0.59002	0.38322	0.02621	0.09602
2	0.90184	0.83734	0.30123	0.01361
3	0.09959	0.55393	0.28458	0.01169
High	0.00629	0.23049	0.68257	0.13833

R^2_{t-1}	Info-Ratio$_{t-1}$ Low	2	3	High
Low	0.34116	0.54684	0.26336	0.33411
2	0.62260	0.50051	0.60799	0.01303
3	0.11183	0.33215	0.03253	0.13248
High	0.01592	0.01451	0.85009	0.03528

Modified Approach

R^2_{t-1}	Alpha$_{t-1}$ Low	2	3	High
Low	0.00409	0.81192	0.40972	0.92009
2	0.03420	0.68527	0.57664	0.14867
3	0.26166	0.02623	0.53094	0.00178
High	0.21068	0.14531	0.05072	0.04242

R^2_{t-1}	Info-Ratio$_{t-1}$ Low	2	3	High
Low	0.00107	0.53752	0.36595	0.39349
2	0.05389	0.57948	0.95157	0.12386
3	0.33594	0.15127	0.35627	0.00490
High	0.43442	0.15715	0.02560	0.02897

Appendix 16 - P-Values from Breusch-Pagan test for all four methodologies for the selectivity predictor
Each portfolio we construct is tested for heteroscedasticity using the Breusch-Pagan test. Values in the matrices are p-values of the test outcome. The test is based on a null hypothesis stating that homoscedasticity is present. Throughout the four approaches, at least four portfolios for each approach reject the null hypothesis, meaning that heteroscedasticity is present. By tendency, this seems to hold especially for the well performing portfolios in the lowest rows. The top portfolio 4:4 shows heteroscedasticity in three out of four approaches. Poorly performing portfolios seem not to be effected by heteroscedasticity.