

UNPACKING PERFORMANCE

A COMPARATIVE STUDY OF VALUE-ADDED AND ALPHA
METRICS IN PREDICTING SWEDISH MUTUAL FUND
MANAGER PERFORMANCE

ANDREAS CHAMIA

TIM ERLANDSSON

Bachelor Thesis

Stockholm School of Economics

2024



Unpacking Performance: A Comparative Study of Value-Added and Alpha Metrics in Predicting Swedish Mutual Fund Manager Performance

Abstract:

This study explores the usage of the gross alpha and value-added metrics to evaluate actively managed mutual fund manager performance in Sweden, inspired by the seminal research of Carhart (1997) and Berk & van Binsbergen (2015). By tailoring and combining their methodologies to the Swedish market, known for its unique investment culture, the paper aims to uncover the predictive power of these metrics in assessing such performance between 2006 and 2019. The empirical findings, derived from a comprehensive dataset from Morningstar Direct, reveal that the value-added measure offers a more precise prediction of mutual fund manager performance as compared to the gross alpha measure. However, the study also highlights the difficulty of definitively attributing fund performance to managerial skill, given the complex interplay of various factors influencing fund returns, coupled with the challenges of data collection inherent in focusing on such a narrowly defined geographical area. These insights contribute to the broader discourse on mutual fund management efficiency, especially within the context of the Swedish investment landscape.

Keywords:

Comparative study, Value-added, Alpha, Mutual fund, Mutual fund manager performance, Benchmark, Fama-French-Carhart

Authors:

Andreas Chamia (25413)
Tim Erlandsson (25410)

Tutors:

Ran Xing, Assistant Professor, Department of Finance at Stockholm University

Examiner:

Adrien d'Avernas, Assistant Professor, Department of Finance

Bachelor Thesis

Bachelor Program in Business and Economics

Stockholm School of Economics

© Andreas Chamia and Tim Erlandsson, 2024



1 Introduction

The subject of mutual fund managerial performance has been a widely debated topic amongst financial academia since the conversation started with Jensen's (1968) highly influential paper. Introducing the so-called Jensen's Alpha, hereafter referred to as the "*alpha*", through the already established capital asset pricing model, or CAPM, Jensen laid the groundwork for a decades-long debate still left unsettled despite numerous scholarly attempts to do so. In our research, numerous contenders have emerged in making such attempts, whether it be Fama's (1970) efficient market hypothesis, Carhart's (1997) paper on the persistence of mutual fund performance, or Berk and van Binsbergen's (2015) paper on measuring skill. Ultimately, despite their various methodologies and approaches, the same lack of consensus remains, as no definitive answers have been produced. At first glance, this debate might seem confined to academia, however, we contend that it also carries practical applications, particularly in the investment sector. This stems from the fact that actively managed mutual funds and their proposed competitor passive index funds make up a lot of the public attention in the area, which subsequently means that settling this debate could potentially have implications for the future of investing.

Considering the above, it is evident that this area of finance is one filled with contention surrounding not only the method of measuring such managerial skill but also its existence in the first place. Furthermore, in conducting the preparatory research for this paper, we have found a persistent geographical focus, namely the United States. This paper aims to combat not only the narrow-minded geographical focus of existing research by focusing on Sweden but also the existing contention within the space. Particularly relevant are the two papers by Carhart (1997) and Berk and van Binsbergen (2015), as the methodology employed in the analysis conducted in our paper is mostly inspired by their respective methodologies, but also due to their contrasting conclusions on the subject matter serving as the major motivation for conducting this study. We propose that by comparing two methods of measurement, namely the alpha metric and the value-added metric as defined by the aforementioned authors, this research paper can move towards an answer to the question of the existence of mutual fund managerial skill and which method is the most accurate in measuring such skill if it exists within the given scope of this paper. Selecting Sweden as the geographical focus of the study also offers a unique opportunity to look at an otherwise underrepresented area of focus given the context, which is simultaneously characterized by factors appealing to such research. These factors include a high level of stock market participation and an early history of mutual fund investing. Additionally, by choosing a singular country instead of a broader scope, such as Europe in its entirety, the dataset benefits from a relatively high level of homogeneity that minimizes the risks of unwanted statistical noise that might stem from a heterogeneous dataset. Sweden is also in a unique position due to its role as a financial hub in the Nordics, its corporate environment and culture, its long history of innovation, the start-up prevalence, its taxation law concerning financial products, concentrated corporate ownership, and the results stemming from the combination of all of these factors. We believe that this creates a unique investing environment that is underutilized as a research focus within the subject, making it an excellent candidate for expanding the available literature.

We have derived three separate hypotheses centered around these set objectives. Our first hypothesis, henceforth also known as *H1*, concerns whether or not mutual fund

managerial skill can be proven to exist. Our belief, and subsequent hypothesis, is that such skill can be observed within our research scope and dataset where the resulting null hypothesis is that such observations cannot be made. In many ways, this first hypothesis constitutes the foundation for the others, and the importance of confirming it should not be understated. The second hypothesis, or $H2$, serves as another crucial piece of the analysis, namely the comparison between the alpha and value-added measurements. We hypothesize that similar to Berk and van Binsbergen (2015) and contrary to Carhart (1997), the value-added measurement will be a better predictor and estimator of mutual fund manager performance. The related null hypothesis to $H2$ is that the alpha measurement will be superior in this regard. Our third and last hypothesis, or $H3$, is related to the second one and examines the subject of whether the benchmark portfolio and the subsequent benchmark factors inspired by Berk and van Binsbergen (2015) are better predictors and estimators of mutual fund manager performance. Our hypothesis is that the benchmark factors will serve better for the intended purpose than the Fama-French-Carhart, or *FFC*, factors to which they are compared, and the subsequent null hypothesis with regard to $H3$ is thus that the *FFC* factors will be superior.

Through analyzing the above-stated three hypotheses, we have confidence that this paper will be able to fulfill the initially set objectives concerning filling the identified gaps in existing literature resulting from, amongst other things, empirical blind spots as a consequence of geographical narrow-mindedness as well as making significant strides in settling the contentious debate around the subject. For instance, when considering the first hypothesis of this paper regarding the existence of managerial skill, Carhart (1997) denied that such a skill could be observed, while Berk and van Binsbergen (2015) came to the opposite conclusion, albeit with a few caveats. Our research thus serves as a valuable addition to the existing literature on the subject, especially given the geographical context. Moreover, outside of academia, there is also some potential for practical results. As mentioned, the available attention span related to investing revolves around the existence of fund managerial skill indirectly, as it entails answering whether to invest in a passively managed index fund versus investing in an actively managed mutual fund. By answering if such a skill exists and, if so, what the best way to measure it is, we believe that this paper has some, albeit geographically limited, potential to affect the future of investing positively.

The aim when designing the desired empirical analysis and the steps involved was to try and replicate the work done in the previously mentioned papers as closely as possible within the context of our comparative study, that is, the comparison of the two chosen performance metrics and their predictive value in evaluating mutual fund manager performance. In doing so, we have utilized their respective methodologies, together with our own original additions, in order to effectively conduct our cross-sectional analysis. The alpha measurement, more specifically the gross alpha, mostly based on Carhart's (1997) work, utilizes a linear regression of monthly Swedish mutual fund returns over either the *FFC* factors as designated in his paper or the benchmark factors as inspired by Berk and van Binsbergen (2015). As for the value-added measurement, it is calculated by utilizing the previously described alphas together with the monthly assets under management, or AUM figure for each mutual fund in the dataset as designated by Berk and van Binsbergen (2015). As a consequence, four sets of monthly values for the individual mutual funds have been acquired, which are then subsequently used to sort the funds into decile portfolios based on their performance according to the respective measurements, those being *FFC* alpha, benchmark alpha, *FFC* value-added, and

benchmark value-added, respectively. The portfolio values for these measurements, as well as the relevant betas and other statistical measurements, are then calculated and compiled for cross-sectional comparison. Concerning the research focus, for reasons that will be further elaborated upon in subsequent sections, this paper will look at Swedish mutual funds between 2006-01-01 and 2019-12-31. The data used consists of four separate datasets: one containing the monthly gross returns of the 461 mutual funds; one containing the AUM values for those same funds, one containing the benchmark fund returns used to construct the benchmark containing the returns of eight index funds, and lastly, one containing the FFC factors in the designated time frame. Acquiring this data involved two different sources of data; for the first three, Morningstar Direct was utilized, while for the latter one, the Swedish House of Finance Research Data Center was used.

Looking at the structure of this paper, it consists of ten parts, of which nine remain after the present introduction and which will be explained in order of appearance. Starting with the *Theoretical Background*, the paper will dive deeper into the existing literature related to our research, as well as the justifications and motivations behind it. In the *Data* section, the datasets used in the analysis will be further examined, including their sources, data selection methods, cleaning procedures, handling of missing values, and other steps taken in their procurement. The *Methodology* section is a closer look at the steps taken to produce our desired results, it is an elaboration on what is described above and includes any potential details necessary for replication. Next, in the *Results* section, we carefully describe the acquired results from the analysis as well as present our tables in a procedural manner to effectively facilitate subsequent discussion and interpretation. In the subsequent *Discussion* section, the results presented in the previous section are interpreted in due course, including the statistical measures, as well as compared where appropriate. Following, in the *Limitations* section, any potential issues in producing the aforementioned results, including data procurement problems, compromises, and any other relevant topics that could have compromised the results and subsequent conclusions stemming from them, are brought up. This sets the stage for the last written section of the paper, namely the *Conclusion* section, in which everything produced thus far is tied together and used in order to come to any potential conclusion with regard to the stated hypotheses. Lastly, in the *References* and *Appendix* sections, we provide our sources in the first, as well as any other relevant information or material used in the latter.

2 Theoretical Background

A The Existence of Skill

The question of which predictor most accurately represents the skill of mutual fund managers derives from the broader question of whether or not mutual fund managers have some inherent skills that they can use to beat the market consistently.

One of the earliest papers from our knowledge that covers the topic of managerial skill is Jensen's (1968) paper on the performance of mutual funds, in which the author investigates the ability of portfolio managers to create greater returns by successfully predicting future stock prices. Jensen finds that on average, actively managed funds were not able to beat a buy-and-hold strategy of the market and that any of the mutual funds that beat the market could not be statistically differentiated from arbitrary luck. Another of the early papers covering this topic was Fama's (1970) paper on the efficient market hypothesis, in which he establishes that stock prices fully reflect all available information in the market. Hence, it would not be possible for actively managed funds to consistently

beat the market by finding mispriced assets. The conclusion Jensen and Fama drew in the late 1960s and early 1970s is later revisited by Carhart (1997) in his paper on the persistence of mutual funds. Carhart measured the alpha for mutual funds and concluded that there was no supporting evidence that skilled or informed mutual fund managers existed that could consistently beat the market over long periods of time. He further emphasizes this point by concluding that if you want to maximize your wealth from investing in mutual funds, you should avoid actively managed funds with higher costs or fees as they have a direct negative impact on performance outcomes. Moreover, these papers also similarly conclude that the performance of actively managed mutual funds is very uncertain and hard to predict.

Conversely, there have also been several studies that suggest that mutual fund managerial skill does exist. Two authors that have together contributed significantly to this topic are Grinblatt and Titman. One of their initial papers, "*Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings*" (1989), established that some mutual funds had significantly higher gross returns than what would be possible if you were simply utilizing a well-known passive investment strategy. Furthermore, in another paper, Grinblatt and Titman (1992) found further evidence of the existence of fund manager skills. By creating a new benchmark free from biases related to firm size, past returns, CAPM betas, and so forth, they established evidence of a difference in performance for mutual funds that is persistent over time, which is consistent with the idea that mutual fund managers have the ability to earn abnormal returns by utilizing their skills for added mutual fund performance. Grinblatt and Titman (1993), also found that some mutual fund managers performed better than the market when you accounted for their portfolio holdings. More specifically, they found that the skills of fund managers who held aggressive growth portfolios could influence the risk-adjusted returns at a statistically significant level. Wermers (2000) is another contributor who found that portfolios held by actively managed mutual funds outperform broad market indices by around 1.3%, of which 70 basis points can be solely attributed to their stock selection skills. More recent contributions include Berk and van Binsbergen's (2015), where the authors measure the skills of a manager by the additional value they can create over a benchmark. Their finding was that the average manager demonstrate skill as they generate, on average, an additional \$3.2 million per year from the capital market over a comparable benchmark portfolio.

In between, there are papers who argue that both sides may be correct; it all depends on the context. One such example is Fama and French (2010), where they tried to determine if abnormal mutual fund returns were based on luck or skill. When looking at mutual fund managers' net returns, they found that they underperformed the returns of a CAPM, three-factor, or four-factor benchmark portfolio. In short, the mutual fund managers needed more skills to cover the costs. Simultaneously, when looking at the results for gross fund returns, they found that there is an existence of managerial skill, both positive and negative.

With that in mind, as mentioned in the introduction, our proposed research builds upon two of the papers above, namely Carhart (1997) and Berk and van Binsbergen (2015), and their respective methodologies and measurements used in evaluating mutual fund managerial skill, which will be further described next.

B *The Alpha Measurement*

The alpha measurement, or more precisely Jensen's alpha, has, since its introduction in Jensen (1968) stood as a staple of the financial industry together with the Capital Asset Pricing Model, or CAPM for short, in evaluating the performance of a given investment portfolio or fund manager. Representing the abnormal return generated by a portfolio in excess of the expected return as defined by the CAPM, in theory, it effectively isolates that of the portfolio's performance, good or bad, which can be attributed to the active management of that particular portfolio. In essence, it allows investors and analysts to gain a nuanced understanding of a fund's value addition, separate from any potential effects of luck or market trends.

However, this is all theoretical, and the measure has its fair share of critics. Additionally, as was claimed by Fama and French (1993), a large amount of the abnormal returns could be explained by other risk factors outside those traditionally part of the model. By adding additional risk factors to the CAPM regression, they argued that a larger amount of the abnormal returns previously attributed to the alpha variable could be explained. These two additional factors that Fama and French found were the book-to-market equity ratio, also commonly referred to as the high minus low factor, or HML, and the size factor, also referred to as small minus big, or SMB. The book-to-market equity ratio reflects what a firm's book value is compared to its common stock. The size factor, relates to a company's market capitalization, which is the per-share price times the outstanding number of shares. When adding these two factors together with the market excess return factor, originally from the CAPM model, a larger amount of what was previously thought of as the alpha could now be explained. In this paper, we will use a combination of the three-factor model and an additional factor called the momentum factor, established by Jegadeesh and Titman (1993), which together create what is known as the Fama-French-Carhart model, or as the FFC model. The momentum factor, or simply MOM, is calculated by taking the average return of the top 30% of stocks minus the returns of the bottom 30% of stocks. By utilizing this model, it is theorized that any resulting alpha is an actual representation of abnormal performance not easily explained by other risk factors, as compared to that calculated through the CAPM.

C *The Value-added Measurement*

Building on this discussion, Berk and van Binsbergen (2015) brings up additional criticism of the alpha measurement in their paper, mainly relating to how it cannot be a valid proxy for skill as, on average, non-zero net alpha investment opportunities must be competed away, which invalidates the measure in its current format for this purpose. Instead, they suggest an alternative measure, the value-added measurement. The value-added measurement encompasses not only the alpha as calculated from, for instance, the FFC model or a benchmark model approach but also the size of the fund, otherwise known as the assets under management figure, as an additional factor in attributing performance. It is argued that the size of a fund is an indication of how well a fund is perceived and, in theory, of how it has performed in the past, as people, if acting rationally, will only buy successful funds. In combining the two measures, they argue that a better, all-encompassing measurement of manager and fund performance has been created. Put differently, funds that perform better over time will draw more capital investments, and as a result, they will have a larger value-added, which will be a reflection of their manager's superior skill. This measure, as described, is what will be used in this paper as the comparative metric to the alpha.

D *Contribution*

Our study aims to contribute to the already existing literature by conducting a comprehensive comparison between the value-added and alpha measurement methods in the context of evaluating a fund manager's performance. As the prior literature has predominantly focused on the US market, our paper will examine Swedish mutual funds, offering an alternative perspective through a previously underutilized geographical research area. In doing so, we hope to bridge the current gap in academia and broaden the available knowledge on mutual fund managerial performance, as well as strengthen future empirical evidence.

3 *Data*

A *Mutual Fund Data*

In conducting the analysis for this paper, a large amount of data pertaining to Swedish mutual funds, index funds, and FFC factor data was required.

The mutual fund data was gathered through the database known as Morningstar Direct. Our extracted dataset contains the necessary information on 461 actively managed Swedish mutual funds needed to conduct our analysis. The relevant data points are the gross monthly return and the monthly assets under management (AUM) of each fund and extend across a time-based scope between 2006-01-01 and 2019-12-31. Moreover, the AUM figures of each fund have been inflation-adjusted to December 2019 based on values from SCB, or Statistics Sweden, as the currency used for those figures was the Swedish krona (SEK).

Arriving at this specific dataset required following a rigorous data selection process, as inspired by Carhart (1997) and Berk and van Binsbergen (2015). More specifically, the mutual funds in the dataset had to be an equity fund, they had to be sold in Sweden, they had to originate from a Swedish company, they had to have a net asset allocation of equity that was greater than 75%, they had to have had an inception date prior to 2017-12-31 to ensure at least 24 months of data, and lastly, they could not be an index fund. The way in which most of these criteria were enforced was through the built-in filtering measures featured on the Morningstar Direct platform. However, due to the nature of our data gathering and the procedures of said platform, one criterion had to be manually enforced, namely that the funds should originate from a Swedish company. This was only applicable when a fund was domiciled outside of Sweden, where the most common such case was funds in the dataset domiciled in Luxembourg while being managed by a Swedish company, for a Swedish market, and so forth. In such cases and similar ones, some additional manual steps had to be taken to ensure that the funds actually originated from a Swedish company. As this was a manual process, we do acknowledge the risk of the dataset containing a small amount of funds that do not adhere to the set criteria, and simultaneously that there might be funds missing that were wrongly categorized. This was a particularly big issue when a fund had been out of existence for a long period of time, however, we deem the potential inaccuracies that this may have caused to be relatively small, if not non-existent, as it was an uncommon issue.

Cleaning and processing the aforementioned dataset was primarily a process of ensuring our two data values, those being monthly gross returns and AUM, could be cross-referenced. Essentially, we needed to ensure that there was a value for each corresponding data point when looking at a specific fund and month, as this was crucial for our regressions. Nevertheless, that caused another issue, namely that the criteria set

out above when initially acquiring the data were in some cases no longer fulfilled, more specifically the ones related to the minimum amount of data for each fund to be included in the dataset. To solve this issue, we enforced a 24-entry per-fund minimum, similar to the demands set by our required inception date from Morningstar Direct. Subsequently, it became a cyclical process of enforcing the minimum requirements set and ensuring simultaneous data points until both were met. Then, for additional statistical accuracy, we chose to enforce the additional criteria of a minimum of 30 entries per month of reported returns for a mutual fund for the first and every subsequent month. The first month that fulfilled these criteria became the starting date for the dataset. At last, we were left with a dataset containing the chosen data points of 461 unique Swedish mutual funds that existed in some capacity between 2006-01-01 and 2019-12-31. Still, issues persisted, as the values reported by Morningstar Direct for AUM contained frequent gaps for many of the mutual funds. In most cases, this is assumed to be due to quarterly reporting. Nonetheless, this means that the data was not complete, and to fill in the missing values, we utilized linear interpolation. This decision was based on many factors, primarily the extent of the calculation, the length of the gaps, and reasonability concerns. The summary statistics of the completed dataset are presented in *Table 1* below.

Table 1: Mutual Fund Summary Statistics

The reported descriptive statistics in this table relate to the dataset containing the 461 mutual funds serving as the basis for this analysis. The value-added and gross alpha value statistics relate to the values for all mutual funds in the dataset. "Assets Under Management" and value-added amounts are in SEK, inflation-adjusted to December 2019 based on data from Statistiska Centralbyrån. The maximum and minimum values of "Monthly Gross Returns" are obtained after filtering out the 5th and 95th percentile values.

Statistic	Mean	Median	Min	Max
Monthly Gross Returns	0.0921%	0.1794%	-74.99%	81.33%
Assets Under Management	3,894,300,528	1,205,425,263	20,839	616,700,000,000
Monthly Value-added from FFC	883,808	12,736	-497,038,450	1,120,972,770
Monthly Value-added from Benchmark	2,431,297	266,389	-708,568,643	693,030,296
Monthly Gross Alpha from FFC	0.0082%	0.0056%	-0.4807%	0.4592%
Monthly Gross Alpha from Benchmark	0.0485%	0.0446%	-0.4919%	0.5301%
Number Of Funds	461			

We are confident that this dataset, having followed the described criteria, will be sufficient to arrive at a result adequate for drawing conclusions around our posed hypotheses. That being said, we are not oblivious to the problems in the data and those that we have come across in acquiring it. For instance, as described, manual processes have been utilized, both in the data selection process and in handling missing values. For the latter, we also acknowledge that by filling in missing values, we have altered the dataset in such a way that it risks not accurately reflecting the true nature of the market during the period. However, as this was an uncommon problem, meaning most funds had complete data within the timeframe, we deem the statistical impact of this to be small. Furthermore, due to what we are assuming to be differing reporting standards as well as a relatively smaller financial market, the data that has been available for this particular study is significantly smaller than that utilized by the papers cited as our inspirations. In other words, compromises had to be made as to the length of the data as well as the funds included in said dataset.

B Benchmark Fund Data

As was done by Berk and van Binsbergen (2015), a benchmark portfolio encompassing the market, in this case, the Swedish market, was needed to conduct our analysis. Through data gathered from Morningstar Direct, in a similar manner to the process utilized when constructing the mutual fund dataset, such a portfolio was constructed, albeit with an

entirely different set of considerations and criteria. However, unlike Berk and van Binsbergen (2015), in choosing to perform our analysis on the Swedish mutual fund market, we also exposed ourselves to a considerably less diverse passively managed index fund market as compared to the one in the U.S. Ultimately, this resulted in a portfolio consisting of 8 index funds, which is smaller than the 11-fund benchmark portfolio utilized in the paper by the above-mentioned authors. The data spans the same timeframe as before, that is between 2006-01-01 and 2019-12-31.

Deriving these 8 index funds required a careful selection process where great considerations were made to as accurately and closely as possible match those market considerations made in the construction of the benchmark portfolio used as inspiration, albeit applied to a smaller and different economy. One such consideration was the popularity of ownership of a given index fund carrying a certain type of market exposure; as described by Berk and van Binsbergen (2015), they were supposed to be so-called market leaders. In making such a determination, data on the popularity of ownership was gathered from two of the largest Swedish stockbrokers, Avanza and Nordnet. It is acknowledged that this data collection is somewhat limited and could lead to arbitrariness in the selection process; however, given other considerations, it was deemed sufficient. Two major impairments emerged during the selection process: for one, the fees of our chosen benchmark portfolio are significantly higher than the ones charged by the Vanguard funds used by Berk and van Binsbergen (2015). For the other, as the Swedish mutual fund market compared to the U.S. is smaller, the exposure to the same market factors is also not as extensive. Another important point to consider is that a large amount of the funds in our benchmark are so-called “index-near” funds, which means that they exclude certain types of investments based on ESG factors. As such, they are not entirely passively following the index; only nearly, but no alternatives were found that still fulfilled other considerations.

In the data cleaning and processing step of constructing the benchmark, a major adjustment to the data had to be made. Looking at *Table 2* below, it can be observed that the inception dates for the funds are considerably different, giving rise to another procedure utilized by Berk and van Binsbergen (2015), namely the orthogonalization process involved in filling out the missing values of the benchmark dataset. In doing so, we utilized QR decomposition to construct an orthogonal basis set based on the initial benchmark dataset, where some slight modifications were made to enable viable results, including replacing the missing values with zeroes. Next, the missing values in the initial benchmark dataset are replaced with the values for the respective funds in the newly constructed orthogonal basis set. After which, the whole process is repeated once more, ultimately creating the complete benchmark factor portfolio. The funds in the benchmark

portfolio, relevant summary statistics, and other relevant information can be seen in *Table 2* below.

Table 2: Benchmark Portfolio Summary Statistics

The reported descriptive statistics in this table relate to the benchmark portfolio of eight index funds. BMF1 consists of European stocks, BMF2 of Swedish stocks, BMF3 of American stocks, BMF4 of Japanese stocks, BMF5 of stocks from emerging markets, BMF6 of global stocks, BMF7 of Asian stocks excluding Japanese stocks, and lastly, BMF8 of stocks of small to medium-sized Swedish companies. The reported values under “Fund Fee” and “Inception Date” were gathered from Morningstar Direct.

Fund	Return Mean	Std Dev	Return Median	Fund Fee	Inception Date
Storebrand Europa A SEK (<i>BMF1</i>)	0.50%	3.74%	0.77%	0.22%	1998-12-30
Storebrand Sverige A SEK (<i>BMF2</i>)	0.84%	4.84%	1.37%	0.21%	1998-12-23
Storebrand USA A SEK (<i>BMF3</i>)	0.86%	3.75%	0.97%	0.21%	1999-10-04
Länsförsäkringar Japan Index (<i>BMF4</i>)	0.31%	4.19%	0.14%	0.22%	1996-01-03
Storebrand Emerging Markets A SEK (<i>BMF5</i>)	0.75%	8.03%	0.42%	0.44%	2010-11-23
Länsförsäkringar Global Index (<i>BMF6</i>)	0.85%	6.58%	0.77%	0.22%	2013-06-11
Swedbank Robur Access Asien A (<i>BMF7</i>)	0.49%	5.36%	0.94%	0.32%	2015-09-23
PLUS Småbolag Sverige Index (<i>BMF8</i>)	1.08%	5.57%	0.42%	0.44%	2017-09-27

It is our firm belief that despite its shortcomings, the benchmark portfolio, although lacking the extensiveness of its inspiration, still encompasses the same broad market considerations and, as such, can be described as sufficient for the analysis in this paper. Simultaneously, we acknowledge that those shortcomings might have affected the accuracy and statistical significance of the related regressions. For one, the higher fees and certain funds being “index-near” rather than strictly following an index presents a significant departure from that which was described in Berk and van Binsbergen (2015), meaning that any potential conclusions come with an asterisk and comparisons are not as easy even notwithstanding the geographical concerns. For the other, the benchmark risks not having the same explanatory value related to the overall market conditions by not being as broad. Furthermore, by filling in the missing values of the dataset, we risk compounding that effect, as the data risks no longer reflecting the true market conditions for the relevant months. Having said that, given the conditions, we deem the constructed benchmark a worthwhile compromise for the analysis.

C Factor Portfolio Data

In enabling our FFC regressions, the factor portfolios constructed by Aytug, Fu, and Sodini (2020), gathered from the Swedish House of Finance Research Data Center website, relating to Sweden and its market conditions between 1983 and 2019, have been utilized. Solely relying on Swedish FFC factor data is a decision motivated by relevance. Given the funds in the dataset, we deem these factors the most applicable. We do, however, acknowledge that this is a compromise given the investment scope of the included mutual funds. The time scope of our analysis spans the period between 2006-01-01 and 2019-12-31, which subsequently means that our dataset of FFC factor portfolio data only contains data for that period.

4 Methodology

A Alpha Measurement

In evaluating the gross alpha measurement, two separate functions for deriving the alpha will be considered, namely the FFC factor model and the benchmark factor model. For the FFC factor model, we will be utilizing the contributions of Carhart (1997), as previously described, to construct the model, as can be seen below in eq. (1):

$$R_{it} = \alpha_{it} + \beta_{it}^{mkt} \text{MKT}_t + \beta_{it}^{sml} \text{SML}_t + \beta_{it}^{hml} \text{HML}_t + \beta_{it}^{mom} \text{MOM}_t \quad (1)$$

Here, *MKT*, *SML*, *HML*, and *MOM* represent the four-factor portfolios of excess return on the market: small minus big, high minus low, and momentum, respectively. Additionally, the beta factors, or β , represent the risk exposures of the i 'th mutual fund, which are estimated through a regression of the fund's monthly returns onto the FFC factors that, as mentioned in *Subsection 3C*, are gathered from Aytug, Fu, and Sodini (2020). Furthermore, the alpha factor, or α , represents the excess return of the funds that cannot be explained by the other risk factors and thus theoretically represents, within the limitations of the regression, the added performance of the funds as a result of the manager's skill. This alpha measure is the *gross alpha* as a result of the inputs into the function being regarded as being prior to adding any fees involved in obtaining it.

The benchmark factor model will follow the same broad template as the one utilized in deriving eq. (1), but instead of the FFC factors, we will utilize our constructed benchmark portfolio as described in *Subsection 3B*, as inspired by Berk and van Binsbergen (2015). The model can be seen below in eq. (2):

$$R_{it} = \alpha_{it} + \beta_{it}^{bm.f1} BMF1_t + \beta_{it}^{bm.f2} BMF2_t + \beta_{it}^{bm.f3} BMF3_t + \beta_{it}^{bm.f4} BMF4_t + \beta_{it}^{bm.f5} BMF5_t + \beta_{it}^{bm.f6} BMF6_t + \beta_{it}^{bm.f7} BMF7_t + \beta_{it}^{bm.f8} BMF8_t \quad (2)$$

In this equation, *BMF1* through *BMF8* represent the monthly returns of the funds in the constructed benchmark portfolio as illustrated by *Table 2*. Similar to eq. (1) the beta factors (β) represent the risk exposures of the i 'th mutual fund, which are estimated through a regression of the fund's monthly returns onto these benchmark fund returns, or factors. Once more, the alpha factor (α) represents the excess return of the funds that cannot be explained by the other risk factors and is explicitly known as the gross alpha measurement.

These models, both eq. (1) and eq. (2), are then utilized to obtain the individual monthly gross alphas for the respective mutual funds in the dataset. That is to say that for each mutual fund, we obtain two separate gross alphas, one based on the FFC factors and another based on the benchmark factors. Carrying out this process involves running a linear regression, containing a rolling window loop with a time window of 18 months, in order to smooth out short-term fluctuations and provide a clearer view of long-term trends. Following, the 461 mutual funds, as described in *Subsection 3A*, are sorted into ten separate decile-ranked portfolios based on the obtained mutual fund monthly gross alphas, coming with a one-month lag. The ranking of the portfolios works such that for each month, the mutual funds having a gross alpha value within the top 10% will be sorted into Portfolio 1, the funds having a gross alpha value within the lowest 10% in Portfolio 10, and so forth. For clarity, as two sets of monthly gross alphas have been obtained, the result of this process is two sets of portfolios based on the gross alphas obtained from their respective models. These portfolios will then be updated monthly throughout our period of examination, that being 2006-01-01 to 2019-12-31. One important thing to note regarding the sorting process is that since the inception dates and termination dates of the mutual funds in our original dataset differ, all mutual funds will not have an alpha value for every month in our dataset. When the portfolios are sorted monthly, it will thus only be the mutual funds with alpha values available for that particular month that will be considered, and not all of the 461 mutual funds.

Having performed the portfolio sorting step of the process, we calculate the monthly gross returns for each portfolio. In addition, a Long-Short portfolio is added to enhance the subsequent analysis, which can be briefly explained as representing an

investment strategy of having a long position in Portfolio 1 and a short position in Portfolio 10. Finally, a linear regression is run based on these portfolios, utilizing eq. (1) and eq. (2), whichever is relevant, out of which a gross alpha is obtained for each of the portfolios, including any relevant statistical measures based on these regressions such as the adjusted R-squared, significance levels, t-statistics, and the betas of the individual factors. The data is then summarized and presented in *Tables 3 and 4* for the FFC model and in *Tables 5 and 6* for the benchmark model.

B Value-added Measurement

Assessing the value-added measurement, as described by Berk and van Binsbergen (2015), involves utilizing the excess gross return of a given fund, see eq. (2) in *Subsection 4A*, the calculated benchmark portfolio return, see eq. (3) below, together with the inflation-adjusted (2019-12) assets under management (AUM) figure, to derive our value-added measurement.

$$R_{it}^B = \sum_{j=1}^{n(t)} \beta_i^j R_t^j \quad (3)$$

In eq. (3), $n(t)$ represents the total number of funds available in the benchmark at time t and β_i^j is the factor found utilizing eq. (2) and running the linear regression, obtaining the individual betas for each of the funds in the benchmark factor portfolio as represented by R_t^j . Utilizing these two equations, we can derive our measurement for value-added as defined in eq. (4):

$$V_{it} = q_{i,t-1} (R_{it}^g - R_{it}^B) \quad (4)$$

Here, the variable q represents the assets under management (AUM) of a given fund one period prior to the current, that is to say, the assets utilized in obtaining the current return figure. In addition, the R_{it}^g variable is the excess gross return of a mutual fund in a given month, as calculated in eq. (2). To clarify, eq. (4), can be described as the product of the benchmark alpha for a given fund and month, together with the corresponding AUM.

However, as we have two separate factor models, and the above-mentioned value-added calculations only apply to the benchmark factor portfolio, we require additional equations. Accordingly, eq. (5) and eq. (6), as seen below, represent the corresponding calculations related to the FFC factor portfolio to obtain a second set of value-added values. In calculating our equivalent to the benchmark return, we calculate the FFC factor portfolio return as seen in eq. (5):

$$R_{it}^{FFC} = \sum_{j=1}^{n(t)} \beta_i^j R_t^j \quad (5)$$

The differences when comparing to eq. (3) is that instead of having betas derived from eq. (2), they are derived from eq. (1), and instead of R_t^j representing the returns of the benchmark factor portfolio funds, it will represent the FFC factors. Subsequently, we can derive our second set of value-added values as defined in eq. (6):

$$V_{it} = q_{i,t-1} (R_{it}^g - R_{it}^{FFC}) \quad (6)$$

Similarly, eq. (6) follows that which was set out in eq. (4), but instead of using R_{it}^B we will be utilizing the values for R_{it}^{FFC} as calculated in eq. (5). Once more, eq. (6) can be

described as the product of the FFC alpha for a given fund and month, together with the corresponding AUM.

Using these two sets of equations, two sets of monthly value-added values for each mutual fund are obtained, one based on the FFC factors, and one based on the benchmark factors. Thereafter, similarly to the alpha-sorted portfolios in *Subsection 4A*, the mutual funds, of which there are 461, are sorted into ten decile-ranked separate portfolios based on these two sets of monthly value-added values. The mutual funds having a value-added value within the top 10% will be sorted into Portfolio 1, the funds having a value-added value within the lowest 10% in Portfolio 10, and so forth, with the result being two sets of portfolios based on their respective models for calculating the value-added measure. Moreover, the additional clarifications related to the two previous sets of portfolios in *Subsection 4A* are still applicable. These portfolios will then be updated monthly throughout our period of examination, that being 2006-01-01 to 2019-12-31.

Similar to the alpha-sorted portfolios, having performed the portfolio sorting step of the process, the gross returns of each portfolio, including the Long-Short portfolio, are then calculated. Finally, a linear regression is run based on these portfolios, utilizing eq. (1) and eq. (2), whichever is relevant, out of which a gross alpha is obtained for each of the portfolios, including any relevant data statistical measures based on these regressions such as the adjusted R-squared, significance levels, t-statistics, and the betas of the individual factors. The data is then summarized and presented in *Tables 7* and *8* for the benchmark model and in *Tables 9* and *10* for the FFC model.

5 Empirical Results

A Preamble

Through the utilization of the methodology described in *Section 4*, a series of four separate sets of portfolios, divided in deciles have been derived, based on either FFC alpha, benchmark alpha, FFC value-added, or benchmark value-added. In the four subsequent sections, a total of eight tables containing the regression results of the performed analysis will be presented. Each respective measure and section are assigned two such tables, where one contains the results stemming from the run regressions on that particular portfolio set, and the other table contains its related statistical measures.

B FFC Alpha Sorted Portfolios

Tables 3 and *4* contain the results stemming from the analysis of the set of portfolios sorted based on the values for the FFC alpha and the related statistical measures, respectively. These outputs are the result of a linear regression where we regress the monthly excess returns for the 10 decile-ranked portfolios in the set on the monthly values

for the four FFC factors, namely being excess market return, or *RMRF*, small-minus-big, or *SMB*, high-minus-low, or *HML*, and the momentum factor, or *MOM*.

Table 3: Regression Results on FFC Alpha Sorted Portfolios

The values presented in this table are the result of a linear regression in which the dependent variable, namely the monthly excess returns of the decile-sorted portfolios consisting of 461 Swedish mutual funds, sorted based on the Fama-French-Carhart (FFC) alpha, are regressed on the monthly values for the four FFC factors, those being the excess market return, or *RMRF*, small-minus-big, or *SMB*, high-minus-low, or *HML*, and the momentum factor, or *MOM*. The standard deviation, or “Std Dev” relates to “Monthly Excess Return”, the “Alpha” measurement is specifically the gross alpha, and the “Long-Short” portfolio is a reflection of the strategy of having a long position in Portfolio 1 and a short position in Portfolio 10. The values for excess return, standard deviation, and alpha are in percentage form, while the rest are in decimal form.

Portfolio	Monthly Excess Return	Std Dev	FFC Factor Model				
			Alpha	RMRF	SMB	HML	MOM
1 (high)	0.0783%	0.4605%	0.0227%	4.85	0.79	0.91	1.16
2	0.0883%	0.4519%	0.0533%	3.81	-0.69	1.20	0.21
3	0.0637%	0.4649%	0.0144%	4.83	1.76	1.01	0.81
4	0.0480%	0.4648%	-0.0019%	4.98	0.66	0.17	0.76
5	0.0753%	0.4962%	0.0445%	5.30	0.83	-1.87	-0.50
6	0.0418%	0.4915%	-0.0159%	5.57	0.91	2.52	0.82
7	0.0639%	0.4505%	0.0148%	5.38	2.92	-0.55	0.70
8	0.0429%	0.4840%	-0.0395%	6.46	1.14	0.53	2.14
9	0.0293%	0.4476%	-0.0091%	4.25	1.59	0.63	0.41
10 (low)	0.0505%	0.4591%	-0.0087%	6.09	1.31	-0.60	0.90
Long-Short	0.0277%	0.4227%	0.0314%	-1.23	-0.51	1.51	0.25

In *Table 3*, several observations can be made. Examining the monthly excess return values for this particular set of portfolios, it is clear that neither their size nor their sign displays any discernible pattern, except for a rough generalization that they are descending in size correlated to the ranking of the portfolios, with Portfolio 2 having the highest and Portfolio 9 the lowest. Furthermore, the values for standard deviation are relatively consistent, roughly hovering around 0.45% to 0.49%, with the Long-Short portfolio being the outlier at 0.42%. Therefore, these values seem to indicate similar levels of volatility across the portfolios except for the Long-Short portfolio being slightly lower, potentially as a result of its self-explanatory hedging characteristics. Moving to the values for the gross alpha, it can be broadly stated that they exhibit the same pattern as the monthly excess return, but interestingly, the alphas exhibit a contextually larger range between the worst-performing and best-performing portfolios. Portfolios 1, 2, 3, 5, and 7, together with the Long-Short portfolio, have positive alphas, suggesting that they outperform relative to the FFC benchmark, with Portfolio 2 performing the best. Portfolios 4, 6, 8, 9, and 10 have negative alphas, implying the opposite. As for the regression coefficients, starting with *RMRF*, it displays significantly larger values for its betas than for the other three, with Portfolio 8 having the largest and Portfolio 2 having the lowest, indicating the highest and lowest exposure to Swedish market conditions, respectively. However, excluding the Long-Short portfolio, which has a negative beta for *RMRF* indicating opposite directionality to the market, they generally hover within the same range. The betas for *SMB* express a significantly larger amount of relative variance within the portfolio set, as well as significantly lower values. Portfolio 7 exhibits the highest beta,

which suggests a strong tilt toward small-cap stocks, whereas, conversely, Portfolio 2 and the Long-Short portfolio have negative betas, suggesting a tilt towards large-cap stocks. The *HML* factor betas are mostly comparable to the *SMB* factor betas when it comes to their range and size; however, here, Portfolio 6 has the highest one, indicating a strong preference toward high book-to-market stocks. On the opposite side, Portfolios 5 and 7 have negative betas, indicating a tilt toward growth stocks. Lastly, looking at the *MOM* factor betas, again, they are roughly comparable to *HML* and *SMB* factor betas when looking at their range and size. Portfolio 8 has the highest beta, implying that the funds within it benefit significantly from following market trends. On the contrary, Portfolio 5 having a negative beta related to this factor could mean that the funds within it follow an investment style that goes against such trends.

Table 4 reports the values for the adjusted R-squared, p-value, and t-value. By putting the regression results in relation to their statistical measures, we can contextualize the results reported in *Table 3*.

Table 4: Statistical Measures of FFC Alpha
Sorted Portfolio Regressions

The reported statistical measures in this table, namely the adjusted R-squared, p-value, and t-values, relate to the alpha derived from the linear regression analysis of the FFC alpha sorted portfolios.

Portfolio	Adj R-squared	p-value	t-value
1 (high)	0.227	0.518	0.648
2	0.172	0.136	1.500
3	0.239	0.682	0.411
4	0.224	0.958	-0.052
5	0.233	0.239	1.183
6	0.329	0.648	-0.457
7	0.305	0.648	0.457
8	0.372	0.236	-1.190
9	0.198	0.793	-0.263
10 (low)	0.340	0.788	-0.270
Long-Short	-0.007	0.394	0.855

The values for adjusted R-squared range from -0.007 for the Long-Short portfolio to 0.372 for Portfolio 8. The Long-Short portfolio is an outlier, with its negative adjusted R-squared suggesting a non-existent predictive value of the model corresponding to it, however, this can likely be attributed to the characteristics of that portfolio and its creation, see *Section 4*. For the other 10 portfolios, the adjusted R-squared hovers around 0.170 to 0.375, indicating a moderate explanatory value of the designated model, but a significant amount of the variance of returns remains unexplained. The p-values makes it clear that none of the portfolios are close to the generally accepted threshold value of less than 0.05. Thus, it becomes difficult to reject the null hypothesis of any of our three stated hypotheses using the results of this particular regression and model. Through the reported t-values, it becomes evident that they vary widely, with positive and negative values displayed. Considering 460 degrees of freedom and a significance level of 5%, the accepted t-value threshold for our one-sample t-test is approximately ± 1.97 , making it clear that none of the portfolios meet such criteria. Nevertheless, the non-significant p-values would negate such a value regardless.

Overall, the reported values for the statistical measures seemingly indicate that the FFC alpha model is not effective at explaining or predicting the performance of

mutual funds, which is further corroborated by the reported values for monthly excess return and alphas in *Table 3* not displaying any discernible pattern in relation to the conducted portfolio ranking procedure.

C *Benchmark Alpha Sorted Portfolios*

Tables 5 and *6* contain the results stemming from the analysis of the set of portfolios sorted based on the values for the benchmark alpha and the related statistical measures, respectively. These tables follow the same general structure as *Tables 3* and *4*, however, exchanging the FFC factors for the benchmark factors. These outputs are the result of a linear regression where we regress the monthly excess returns for the 10 portfolios in the set on the monthly values for the 8 benchmark funds, namely, *BMF1*, *BMF2*, *BMF3*, *BMF4*, *BMF5*, *BMF6*, *BMF7*, and *BMF8*.

Table 5: Regression Results on Benchmark Alpha Sorted Portfolios

The values presented in this table are the result of a linear regression in which the dependent variable, namely the monthly excess returns of the decile-sorted portfolios consisting of 461 Swedish mutual funds, sorted based on the benchmark alpha, are regressed on the monthly values for the 8 index funds in the benchmark portfolio, namely, *BMF1*, *BMF2*, *BMF3*, *BMF4*, *BMF5*, *BMF6*, *BMF7*, and *BMF8*. The standard deviation, or “Std Dev” relates to “Monthly Excess Return”, the “Alpha” measurement is specifically the gross alpha, and the “Long-Short” portfolio is a reflection of the strategy of having a long position in Portfolio 1 and a short position in Portfolio 10. The values for excess return, standard deviation, and alpha are in percentage form, while the rest are in decimal form.

Portfolio	Monthly Excess Return	Std Dev	Benchmark Factor Model								
			Alpha	BMF1	BMF2	BMF3	BMF4	BMF5	BMF6	BMF7	BMF8
1 (high)	0.0786%	0.4889%	0.0693%	-0.20	4.78	3.58	-1.42	-0.29	-1.45	0.25	-0.50
2	0.0870%	0.4851%	0.0571%	-1.27	2.17	5.28	-3.20	0.25	0.04	1.24	0.51
3	0.0804%	0.4724%	0.0405%	0.04	2.33	3.80	-1.68	-0.73	-0.15	0.50	1.42
4	0.0859%	0.4827%	0.0666%	-0.45	1.63	5.69	-3.34	0.45	-0.84	-0.10	1.12
5	0.0523%	0.4783%	0.0295%	-0.28	2.43	4.53	-1.39	-0.29	-0.94	-0.33	0.71
6	0.0895%	0.4633%	0.0645%	-0.69	6.69	2.24	-2.61	-0.14	-0.32	0.24	0.78
7	0.0861%	0.4542%	0.0546%	-0.71	2.26	5.40	-3.27	-0.21	-0.10	0.33	1.24
8	0.0889%	0.4795%	0.0558%	-2.09	4.00	4.75	-1.15	-0.28	-0.65	-0.09	0.86
9	0.0018%	0.4654%	-0.0255%	-0.78	3.68	4.83	-2.52	-1.56	-0.82	-0.22	1.62
10 (low)	0.0692%	0.4717%	0.0417%	-0.84	3.86	4.46	-2.41	-0.63	-0.46	0.01	0.98
Long-Short	0.0095%	0.4818%	0.0276%	0.64	0.92	-0.89	0.99	0.34	-0.99	0.25	-1.48

Looking at *Table 5* and the values for monthly excess return, it is observable that they are fairly consistent across the 10 portfolios, generally ranging from 0.0523% to 0.0895%. The noticeable outlier is Portfolio 9, which significantly underperforms the other portfolios at 0.0018%. In addition, the range between the best-performing and worst-performing portfolios is contextually rather large. As for the standard deviations, it can be observed that they are relatively uniform across all the portfolios, ranging from around 0.45% to 0.49%, suggesting that the portfolios have similar levels of risk exposure. Through the reported gross alphas, it can be stated, although not clear and consistent, that there is some kind of relationship with regard to the portfolio rankings, with Portfolio 1 performing the best. However, moving down the list, the pattern becomes less discernible. It is worth noting that all portfolios except for Portfolio 9 display positive alphas, indicating outperformance relative to the benchmark, and also that the range of values between best-performing and worst-performing is comparable to that for the monthly excess return values. Turning to the betas for the regression coefficients and referring back to *Table 2*, several observations can be made. Starting with *BMF1*, which represents an index fund with a European focus, the betas are mostly negative or only slightly positive, indicating an inverse relationship to that specific market exposure. Conversely, the betas for *BMF2* and *BMF3*, representing exposure to Swedish stocks and American stocks, respectively, are noticeably high across the portfolios, signaling a significant exposure to such market factors. The betas for *BMF4* through *BMF6* are mostly negative, with *BMF4* being the most extreme, suggesting a significant aversion to Japanese stocks.

BMF7 is a mixed set of positive and negative betas, and *BMF8* is a similar story, albeit with fewer negative ones, but neither has a discernible pattern. This suggests that exposure to Asian stocks, excluding Japanese ones, and to Swedish small-to-medium-cap stocks, respectively, has no clear relationship with fund performance. Other specific patterns or correlations are hard to come by, as the patterns exhibited by the betas are not clearly correlated to the reported alphas outside those observations already stated. A noticeable exclusion in the above analysis is the Long-Short portfolio, but as it builds on the exposures of Portfolios 1 and 10, such an analysis is redundant, thus, we refer to those portfolios on which it was based.

Compared to the results reported in *Subsection 5B* for the FFC alpha regressions, the reported monthly excess returns are mostly larger and exhibit a somewhat clearer pattern. This is similar to the reported alphas also displaying such trends. The standard deviations are mostly comparable, both in range and pattern. These results could indicate that the benchmark factors are slightly better than the FFC factors as a predictor within the context.

Table 6 reports the values for the adjusted R-squared, p-value, and t-value. By putting the regression results in relation to their statistical measures, we can contextualize the results reported in *Table 5*.

Table 6: Statistical Measures of Benchmark Alpha
Sorted Portfolio Regressions

The reported statistical measures in this table, namely the adjusted R-squared, p-value, and t-values, relate to the alpha derived from the linear regression analysis of the benchmark alpha sorted portfolios.

Portfolio	Adj R-squared	p-value	t-value
1 (high)	0.263	0.061	1.886
2	0.299	0.111	1.604
3	0.225	0.268	1.113
4	0.311	0.060	1.899
5	0.261	0.413	0.820
6	0.354	0.050	1.978
7	0.361	0.088	1.718
8	0.395	0.090	1.709
9	0.329	0.447	-0.763
10 (low)	0.317	0.224	1.221
Long-Short	-0.005	0.515	0.652

The values for adjusted R-squared range between -0.005 for the Long-Short portfolio and 0.395 for Portfolio 8. Once more, similar to the FFC alpha regression, the Long-Short portfolio is an outlier, having a negative adjusted R-squared, but this is to be expected. For the other 10 portfolios, these values hover around 0.225 to 0.395, which, compared to the results reported in the previous section, is slightly higher, though overall, it still indicates a moderate explanatory value. Observing the reported p-values, although trending lower than those reported for the FFC alpha regression, here, one portfolio clears the designated threshold of less than 0.05, namely Portfolio 6. Though, similar to before, it is difficult to reject the null hypothesis of our three stated hypotheses using the results of this particular regression and model. In reference to the reported t-values, it is observable that they vary widely, but only one clears the previously calculated t-value threshold of ± 1.97 , namely Portfolio 6, which is also the portfolio displaying significant results.

On the whole, outside of significant results for a singular portfolio, the reported values for the statistical measures seemingly indicate that the benchmark alpha model is not effective at explaining or predicting the performance of mutual funds, which is similar to that stated about the FFC alpha model. However, unlike the FFC alpha model, the reported values for monthly excess return and alphas, as seen in *Table 5*, exhibit a somewhat more discernible pattern as related to the portfolio ranking procedure.

D Benchmark Value-added Sorted Portfolios

Tables 7 and *8* contain the results stemming from the analysis of the set of portfolios sorted based on the values for the benchmark value-added and the related statistical measures, respectively. These tables follow the same structure as *Tables 5* and *6*, but add a value-added variable. These outputs are the result of a linear regression where we regress the monthly excess returns for the 10 portfolios in the set on the monthly values for the 8 benchmark funds, namely, *BMF1*, *BMF2*, *BMF3*, *BMF4*, *BMF5*, *BMF6*, *BMF7*, and *BMF8*.

Table 7: Regression Results on Benchmark Value-Added Sorted Portfolios

The values presented in this table are the result of a linear regression in which the dependent variable, namely the monthly excess returns of the decile-sorted portfolios consisting of 461 Swedish mutual funds, sorted based on the benchmark value-added, are regressed on the monthly values for the 8 index funds in the benchmark portfolio, namely, *BMF1*, *BMF2*, *BMF3*, *BMF4*, *BMF5*, *BMF6*, *BMF7*, and *BMF8*. The standard deviation, or “Std Dev” relates to “Monthly Excess Return”, the “Alpha” measurement is specifically the gross alpha, the value-added is derived from a function utilizing the derived benchmark alphas with the assets under management figures in the mutual fund dataset, and the “Long-Short” portfolio is a reflection of the strategy of having a long position in Portfolio 1 and a short position in Portfolio 10. The values for excess return, standard deviation, and alpha are in percentage form; the values for value-added are in SEK inflation-adjusted to December 2019 based on data from Statistiska Centralbyrån; while the rest are in decimal form.

Portfolio	Monthly Excess Return	Std Dev	Value-added	Benchmark Factor Model								
				Alpha	BMF1	BMF2	BMF3	BMF4	BMF5	BMF6	BMF7	BMF8
1 (high)	0.1642%	0.4519%	23,096,466	0.1417%	-1.81	2.83	4.40	-3.45	0.62	-0.53	-0.17	1.99
2	0.1296%	0.4869%	5,461,465	0.1030%	-1.26	1.66	6.59	-3.34	-0.20	-0.84	-0.06	1.60
3	0.0977%	0.4976%	2,562,862	0.0935%	-0.80	4.16	3.26	-4.01	1.38	-0.98	0.02	1.00
4	0.0789%	0.4779%	1,310,606	0.0537%	0.60	1.53	6.18	-4.65	-0.51	-0.13	0.13	1.30
5	0.0884%	0.4756%	652,231	0.0610%	-1.43	5.41	3.54	-2.79	-0.58	-0.46	0.20	1.37
6	0.0870%	0.4334%	276,164	0.0832%	-1.02	4.69	3.74	-3.67	-0.23	-1.01	-0.10	0.81
7	0.0496%	0.4481%	42,243	0.0197%	-2.25	2.41	2.79	-0.32	-0.27	0.30	-0.14	0.84
8	0.0262%	0.4861%	-238,177	0.0004%	-0.66	2.49	4.98	-1.79	-0.11	-0.66	-0.12	0.70
9	0.0013%	0.4532%	-891,248	-0.0295%	0.21	4.38	4.02	-2.30	-0.53	-0.42	0.32	0.66
10 (low)	-0.0214%	0.4978%	-7,082,294	-0.0368%	0.39	3.99	5.38	-5.73	-0.56	-0.34	0.08	1.30
Long-Short	0.1856%	0.4394%	30,178,760	0.1785%	-2.20	-1.16	-0.98	2.28	1.18	-0.19	-0.24	0.69

Referencing *Table 7* and its reported values for monthly excess return, it is noticeable, dissimilar from the models in the two previous sections, that they exhibit a significantly clearer pattern with regard to the rankings of the portfolios. In exemplifying, we see that Portfolio 1 has the largest return of the ranked portfolios at 0.1642%, and Portfolio 10 has the lowest value by a margin at -0.0214%. Continuing the comparison, both the range and size of these values are also larger. The standard deviations, however, exhibit a similar pattern as the previous models, albeit with a slightly larger range, ranging from around 0.43% to 0.50%, but generally, they can still be interpreted as the portfolios having a similar level of risk exposure. As for the reported gross alphas, what could be said about the monthly excess return is still applicable in all respects. Portfolio 1 displays the highest alpha of the ranked portfolios at 0.1417%, and Portfolio 10 is the lowest at -0.0368%. This brings us to the reported values for value-added, which, once more, exhibit such a pattern. Portfolio 1 performs the best of the ranked portfolios at 23,096,466 SEK, and Portfolio 10 performs the worst at -7,082,294 SEK. Focusing on the Long-Short portfolio, it can be stated that in all three aspects, it performs the best, which seems to solidify that, at least superficially, the benchmark value-added model performs the best in predicting performance thus far. Shifting emphasis towards the betas of the regression coefficients and referring back to *Table 2*, it can be roughly stated that what was observed with respect to the benchmark alpha model reported in *Table 5* remains applicable here. Factor *BMF1*, that being the exposure to European stocks, has betas that are mostly negative with a few

outliers. Noticeably, Portfolio 4 has a beta of 0.6, indicating a larger exposure to this market than the other portfolios. Similar to the results reported for the benchmark alpha, the betas for *BMF2* and *BMF3*, representing exposure to Swedish stocks and American stocks, respectively, remain within the same size range; however, the relative size of the individual portfolio betas has changed significantly. Furthermore, the betas for *BMF4* through *BMF6* remain mostly negative, with *BMF4* being the most extreme, suggesting a significant aversion to Japanese stocks here as well. *BMF7* is still a mixed set of positive and negative betas; however, *BMF8* is where it starts to differ, as the betas are now positive across the portfolios. This means that, although exposure to Asian stocks, excluding Japanese ones, remains the same, exposure to Swedish small-to-medium-cap stocks has turned positive. In addition, it can be observed that the individual factor betas for the different portfolios have changed, and in some cases significantly so. This, coupled with the reported performance values, seems to indicate that the benchmark value-added model is better at predicting what types of market exposures are successful in the long run and the subsequent performance benefits as compared to the previous two models, but it should be noted that no such relationship can be definitely proven.

Considering the above, when compared to the models utilized in *Subsections 5B* and *5C*, it is reasonable to argue that the benchmark value-added model seems to be a better predictor of performance than the other two such models.

Table 8 reports the values for the adjusted R-squared, p-value, and t-value. By putting the regression results in relation to their statistical measures, we can contextualize the results reported in *Table 7*.

Table 8: Statistical Measures of Benchmark Value-added Sorted Portfolio Regressions

The reported statistical measures in this table, namely the adjusted R-squared, p-value, and t-values, relate to the alpha derived from the linear regression analysis of the benchmark value-added sorted portfolios.

Portfolio	Adj R-squared	p-value	t-value
1 (high)	0.356	0.000	4.469
2	0.378	0.003	3.068
3	0.245	0.015	2.473
4	0.325	0.120	1.563
5	0.335	0.074	1.798
6	0.314	0.009	2.650
7	0.169	0.582	0.552
8	0.288	0.991	0.011
9	0.374	0.349	-0.941
10 (low)	0.341	0.300	-1.040
Long-Short	-0.003	0.000	4.638

Directing our attention to the reported values for the adjusted R-squared, we can see that they range between -0.003 for the Long-Short portfolio and 0.378 for Portfolio 2. Again, similar to the previous sections, the Long-Short portfolio is as expected an outlier, having a negative adjusted R-squared. The values for the other 10 portfolios hover around 0.169 to 0.378, which is lower than those for the benchmark alpha but generally higher than those of the FFC alpha. Still, it indicates a moderate explanatory value. When observing the reported p-values, a difference emerges. For the benchmark value-added model, a significantly larger amount of the portfolios exhibit acceptable levels of significance at lower than 0.05 compared with the previous models. Notably, Portfolio 1, a crucial

component of the analysis, has a significance level of approximately 0. Still, the small majority, still exhibit levels of insignificance. Turning toward the p-values, they still vary widely, but for the portfolios exhibiting acceptable significance levels, namely portfolios 1, 2, 3, 6, and the Long-Short portfolio, they all display t-values clearing the previously calculated t-value threshold of ± 1.97 .

Altogether, the statistical measures seemingly provide a relatively stable ground on which a continued analysis is possible. Additionally, in contrast to the models in *Subsections 5B* and *5C*, the reported values for monthly excess return and gross alphas, as seen in *Table 7*, are entirely in line with the portfolio ranking procedure. In other words, by looking at the results of both *Tables 7* and *8* and considering their implications, it is reasonable to conclude that, unlike the previous models, the benchmark value-added model seems to be significantly more effective in explaining or predicting the performance of the mutual funds within our dataset. Thus, rejecting the null hypothesis of our three stated hypotheses using these results seems significantly more realistic.

E FFC Value-added Sorted Portfolios

Tables 9 and *10* contain the results stemming from the analysis of the set of portfolios sorted based on the values for FFC value-added and the related statistical measures, respectively. These tables follow the same structure as *Tables 3* and *4*, but add a value-added variable. These outputs are the result of a linear regression where we regress the monthly excess returns for the 10 portfolios in the set on the monthly values for the four FFC factors, namely the excess market return, or *RMRF*, small-minus-big, or *SMB*, high-minus-low, or *HML*, and the momentum factor, or *MOM*.

Table 9: Regression Results on FFC Value-added Sorted Portfolios

The values presented in this table are the result of a linear regression in which the dependent variable, namely the monthly excess returns of the decile-sorted portfolios consisting of 461 Swedish mutual funds, sorted based on the Fama-French-Carhart (FFC) value-added, are regressed on the monthly values for the four FFC factors, those being the excess market return, or *RMRF*, small-minus-big, or *SMB*, high-minus-low, or *HML*, and the momentum factor, or *MOM*. The standard deviation, or “Std Dev” relates to “Monthly Excess Return”, the “Alpha” measurement is specifically the gross alpha, the value-added is derived from a function utilizing the derived FFC alphas with the assets under management figures in the mutual fund dataset, and the “Long-Short” portfolio is a reflection of the strategy of having a long position in Portfolio 1 and a short position in Portfolio 10. The values for excess return, standard deviation, and alpha are in percentage form; the values for value-added are in SEK inflation-adjusted to December 2019 based on data from Statistiska Centralbyrån; while the rest are in decimal form.

Portfolio	Monthly Excess Return	Std Dev	Value-added	FFC Factor Model				
				Alpha	RMRF	SMB	HML	MOM
1 (high)	0.1660%	0.4464%	22,066,987	0.1140%	4.83	0.84	0.43	1.02
2	0.1406%	0.4665%	3,579,425	0.0984%	5.18	0.80	0.36	0.21
3	0.0855%	0.4941%	1,427,585	0.0448%	4.62	0.55	0.87	0.33
4	0.0809%	0.4888%	563,137	0.0523%	4.37	1.31	-0.01	-0.21
5	0.0790%	0.4294%	133,274	0.0523%	3.62	1.46	1.64	-0.08
6	0.0605%	0.4492%	-156,739	0.0253%	3.75	0.50	2.18	0.31
7	0.0613%	0.4762%	-523,014	0.0243%	4.69	2.88	0.48	0.24
8	0.0505%	0.4853%	-1,250,686	-0.0102%	6.04	-0.62	-0.24	0.93
9	-0.0202%	0.4855%	-2,773,968	-0.0881%	5.87	-0.13	0.42	1.47
10 (low)	-0.0554%	0.5116%	-13,237,819	-0.1229%	5.97	1.33	-0.16	1.54
Long-Short	0.2213%	0.4997%	35,304,806	0.2368%	-1.14	-0.49	0.60	-0.52

In reference to *Table 9* and the reported values for monthly excess return, it can be said, similar to the values related to the benchmark value-added model, that they display a clearly discernible pattern with regard to the portfolio rankings. Moreover, the range and size of these values are also larger, with Portfolio 1 having the highest return of the individual portfolios at 0.1660% and Portfolio 10 having the lowest at -0.0554%. The

standard deviations exhibit a pattern similar to the previous three models, ranging between roughly 0.42% and 0.51%, which can once again be interpreted as the portfolios having a similar level of risk exposure. In terms of the gross alphas, in comparison to *Subsection 5D* and *Table 7*, it can be stated that the range of these values is higher, but the peak is lower, with the highest alpha at 0.1140% for Portfolio 1 and the lowest alpha at -0.1229% for Portfolio 10, when considering the ranked portfolios. The values for the value-added exhibit the same pattern, having a peak with Portfolio 1 at 22,066,987 SEK for the ranked portfolios, with the corresponding lowest value attributed to Portfolio 10 at -13,237,819 SEK. Interestingly, the value-added measure seems to have an even split of positive and negative values going directly down the middle of the portfolio set. Pivoting towards the Long-Short portfolio, similar to that which was the case for the benchmark value-added model, in all three aspects, it performs the best, even more than the aforementioned model as a result of the wider range between best- and worst-performing portfolios. Seemingly, this seems to support the other indications that the FFC value-added is also relatively accurate when predicting performance. Transitioning to the regression coefficients, the initial impression is that the reported betas for the individual factors display a generally comparable behavior to that which was reported in *Table 3*. The *RMRF* betas are significantly larger than those for the other three, with Portfolio 8 having the largest value and Portfolio 5 having the lowest value out of the ranked portfolios, indicating the highest and lowest exposure to market conditions, respectively. However, excluding the Long-Short portfolio, which has a negative beta for *RMRF* indicating opposite directionality to the market, they typically hover within the same range, which was also the case for the aforementioned model. Following, the betas for the *SMB* factor are significantly smaller. Portfolio 7 exhibits the highest beta, which suggests a strong tilt toward small-cap stocks, whereas, conversely, portfolios 8, 9, and the Long-Short portfolio have negative betas, suggesting a tilt towards large-cap stocks. As for the *HML* factor betas, their size and range are generally comparable to those of the *SMB* factor betas, with Portfolio 6 having the largest value at 2.18, indicating a strong preference toward high book-to-market stocks, and Portfolio 8 having the lowest value at -0.24, indicating a tilt toward growth stocks. Proceeding to the *MOM* factor betas, again, they are roughly comparable to the *HML* and *SMB* factor betas when looking at their range and size, with Portfolio 10 having the highest beta at 1.54, implying that funds within it benefit significantly from following market trends, and Portfolio 4 having the lowest beta out of the ranked portfolios at -0.21, implying that the funds within it follow an investment style that goes against such trends.

By comparing these results with those stemming from the benchmark value-added model in *Subsection 5D* and the two earlier models presented in *Subsections 5B* and *5C*, it is reasonable to argue that the FFC value-added model seems to be a better predictor of performance than these two previously presented models when considering what has been presented above. However, when comparing this model to the benchmark value-added model, a clear interpretation is harder to achieve without also considering

the related statistical measures and putting the two aspects in relation to each other; therefore, *Table 10* will be looked at next.

Table 10: Statistical Measures of FFC Value-added
Sorted Portfolio Regressions

The reported statistical measures in this table, namely the adjusted R-squared, p-value, and t-values, relate to the alpha derived from the linear regression analysis of the FFC value-added sorted portfolios.

Portfolio	Adj R-squared	p-value	t-value
1 (high)	0.230	0.000	3.376
2	0.268	0.005	2.859
3	0.191	0.245	1.168
4	0.181	0.173	1.371
5	0.216	0.113	1.596
6	0.193	0.469	0.726
7	0.242	0.497	0.681
8	0.307	0.770	-0.293
9	0.293	0.013	-2.504
10 (low)	0.262	0.001	-3.242
Long-Short	-0.014	0.000	5.461

The values for the adjusted R-squared range from -0.014 for the Long-Short portfolio to 0.307 for Portfolio 8. Once again, similar to the previous models, the Long-Short portfolio is an expected outlier, having a negative adjusted R-squared. The values for the other 10 portfolios hover around 0.181 to 0.307, which is the lowest out of all the presented models. However, despite this model having the lowest adjusted R-squared values, they still indicate a moderate explanatory value. Shifting focus to the reported p-values, they tend to be lower than those reported for the benchmark value-added regression; however, the number of significant values remains the same. Important to note, however, is that in addition to Portfolio 1 having a significant value of 0, which was also the case in the aforementioned regression, Portfolio 10 also has a significant value of 0.001. This is remarkable because it means that the two arguably most critical results are significant. Looking at the t-values further corroborates this fact, as all significant results also exhibit t-values clearing the previously calculated t-value threshold of ± 1.97 , with the Long-Short portfolio having the highest value at 5.461. Though it should be noted that the majority of the portfolios still exhibit some level of insignificance.

Considering all the above, the statistical measures seemingly provide a relatively stable ground on which a continued analysis is possible. Furthermore, as was the case with the benchmark value-added model, the reported values for monthly excess return and gross alphas, as seen in *Table 9*, are entirely in line with the portfolio ranking procedure. Put differently, by considering the implications of *Tables 9* and *10*, it is reasonable to conclude that the FFC value-added model is at least equally effective as the benchmark value-added model in explaining or predicting the performance of the mutual funds within our dataset. As such, rejecting the null hypothesis of our three hypotheses using the results of this particular regression seems realistic.

6 Discussion

A Interpretation of Results

In the preceding section, it was implied that multiple interpretations could be made by observing the data gathered from our four respective models, as presented in *Tables 3, 5, 7, and 9*.

Comparing the monthly excess returns of the models, it is evident that a difference exists between the results of the alpha-sorted portfolios and the value-added sorted portfolios pertaining to the recognition of any patterns related to their returns. Furthermore, in looking at the return spread of the alpha-sorted portfolios, it can be observed that the difference between the highest and lowest return is only 0.0590% for the FFC version and 0.0877% for the version using the benchmark factors. On the contrary, when observing the portfolios sorted by the value-added measure, they have a spread of 0.2214% for the FFC version and 0.1856% when utilizing the benchmark factors. A wider spread could indicate that there is some sort of diversification when it comes to the performance between the portfolios, and judging by the results we found, the value-added sorting method seemingly captures it most effectively. This can be corroborated by looking at the Long-Short portfolios of each version. To elaborate, since the Long-Short portfolios are, put simply, the difference between portfolios 1 and 10 for a given metric, the higher the value exhibited by the Long-Short portfolio, the better the related model is at capturing and predicting the spread in the performance between the best-performing and worst-performing mutual funds with regard to said metric. Taking the excess monthly return of the four models and ranking them with regard to this predictability power from highest to lowest, FFC value-added comes first, followed by benchmark value-added, FFC alpha, and finally, benchmark alpha.

When comparing the gross alphas between our four models, a similar pattern emerges to the one displayed by the monthly excess returns. Surveying the Long-Short alphas for the four models, we find that the FFC value-added has an alpha of 0.2368%, the benchmark value-added has an alpha of 0.1785%, the FFC alpha has an alpha of 0.0314%, and finally, the benchmark alpha has the lowest alpha value of 0.0276%. In this case, since both of the value-added models outperform their alpha model counterparts, as well as display a more discernible pattern related to their respective portfolios and rankings, it could be argued that the value-added models have a greater predictive ability with regard to the alpha. Additionally, comparing the FFC versions with the benchmark versions of our models, the FFC versions generally outperform the benchmark versions when it comes to this ability. Though it should be noted that the difference between the two is comparatively small when contrasted with the difference between the value-added and alpha models.

In analyzing the value-added values between the FFC and benchmark model, all portfolios down until Portfolio 8 produce a positive value-added in the benchmark model, whereas, conversely, in the FFC models, the last positive value-added portfolio is Portfolio 5 before turning negative. When performing the Long-Short portfolio analysis, it is reasonable to assert that the FFC factors outperform the benchmark factors when it comes to the predictive capabilities with regard to the value-added metric. An interesting thing to note is that the positive value-added numbers for both related models are larger than the negative ones, so on an aggregate level, mutual fund managers seem to be adding value to the market compared with a passive alternative. One explanation as to why the overall value-added is positive is the tendency of investors to allocate more resources

toward better-performing mutual funds, or, put differently, that better-performing mutual funds attract more capital on average. If top-performing mutual funds can consistently attract more capital and resources while still delivering positive returns, it could be an indication of the existence of managerial skills. However, it is important to note that these values do not include any potential fees charged by the mutual funds for the purposes of this analysis; therefore, the value-added that makes its way to investors is not entirely known, and as such, any potential practical implications are limited.

B *Interpretation of Statistical Measures*

Referring back to *Tables 4, 6, 8, and 10*, it is evident that the statistical measures of the four different models of performance measurement varied widely in their effectiveness, at least as interpreted by the measures reported in these four tables.

When looking at the values for adjusted R-squared, it can be observed that, although figures differ between the models, in general, they can all be categorized as indicating that the models have a moderate explanatory value with respect to the performance of the funds in the dataset. However, it should be noted that the models utilizing the benchmark factors, exhibit observably higher explanatory values than those showcased by the models utilizing the FFC factors. Furthermore, the adjusted R-squared values resulting from the benchmark value-added regression, are the lowest ones. One interpretation of the above is that, according to this specific measurement, the benchmark factors seemingly better capture the market exposures that affect the performance of the mutual funds present in the analysis. This is a natural effect of the plurality involved when comparing the FFC factors versus the benchmark factors. As there are more benchmark factors, encompassing a theoretically wider net of market considerations, one can imagine that models utilizing them should exhibit a higher explanatory value. It is important to note that although the effect is easily observed, it is not sufficiently large to change the moderate explanatory value of the models. Furthermore, one should take into consideration the difficulties involved when performing the kind of analysis performed in this paper; constructing models explaining the performance of mutual funds is notably difficult, so the achieved levels of explanatory values are deemed sufficient for the purposes of this analysis. The moderate explanatory values of the four models could also be construed as an indication that there are other factors at play in the performance of the mutual funds in the dataset, which, among other things, could be an indication of the existence of managerial skill as such factors are absent from the constructed models. An additional note is that the small, slightly negative adjusted R-squared as expressed for the Long-Short portfolios of all four models can most likely be explained by the fact that the designated linear model has a hard time explaining the combined effect of the large positive alphas of Portfolio 1 and the low and often negative alphas of Portfolio 10. The model is not equipped to explain the combined results of the two portfolios with respect to the benchmark factors or the FFC factors. Still, it should be noted that there are most likely other such effects as well, this was simply deemed the major cause. That being said, it is a natural effect of the construction of such a portfolio and therefore not an immediate cause of concern in the context of the analysis and conclusions drawn from it.

Following the above, in examining the p-values for the four models, an immediate observation is that for the regressions related to the alpha models, whether based on benchmark factors or FFC factors, they are mostly insignificant. Although it should be noted that the values for the benchmark alpha model are noticeably closer to significance in general, as well as there being one significant value. In the broader context, however,

that is unimportant as the levels of insignificance in both related models lead to it being hard to draw any concrete conclusion from their results as the relationship might be due to random chance without there being any real pattern of effect between the variables. Conversely, when looking at the value-added models, whether based on benchmark factors or FFC factors, their reported values exhibit a significantly higher level of significance. Both versions show significance for five of the portfolios in their respective portfolio sets when including the Long-Short portfolio. Furthermore, both models exhibit an exceptionally high significance level of 0.00 when looking at the results of Portfolio 1, a crucial part of the conducted analysis. However, the benchmark value-added model performs significantly worse in the lower half of the portfolio set, having only insignificant values from Portfolio 7 onwards. Notably, in the FFC value-added model, Portfolio 10, another crucial part of the analysis, exhibits a significant value of 0.01, which puts it ahead of the benchmark value-added model. It should also be noted that for both the benchmark value-added model and the FFC value-added model, the Long-Short portfolio exhibits significance levels of 0.00. All in all, the reported p-values for these latter two models make it a significantly more reasonable possibility to draw tangible conclusions with respect to the posed hypotheses, as there is a 5% or less risk of making a Type I error when analyzing the applicable values.

Lastly, by considering the t-values for the four models, it can be stated that roughly the same conclusions as those drawn in the previous paragraph are still applicable. Looking at the alpha models, the majority of the portfolios in both sets have t-values that are not clearing the previously calculated threshold of ± 1.97 . The t-value for Portfolio 6 out of the benchmark alpha-sorted portfolios is the one exception, which is a natural effect of that being the only significant value out of these two models. When looking at the value-added models, it can be observed that all those portfolios that have been identified as having significant results have t-values clearing this threshold of ± 1.97 . This essentially indicates that those values identified as significant are also likely due to that which is incorporated into the respective models, rather than random chance. Thus, in combination with the results of the previous two paragraphs, drawing conclusions with respect to the three hypotheses posed in the introduction is seemingly possible.

C *Supplementary Interpretations*

In conducting a deeper analysis of the beta values for the benchmark models, as seen in *Tables 5 and 7*, the mutual funds in our dataset seem to have a greater exposure towards *BMF2* and *BMF3*, which are to Swedish and American stocks, respectively. An explanation for the higher exposure to Swedish stocks can be attributed to our selection of mutual fund data, where we specifically decided to focus on the Swedish mutual fund market. Even though Swedish mutual funds are able to invest globally, there seems to be an overexposure to Swedish stocks, which a familiarity bias could partially explain. It could also be explained by the fact that Swedish mutual fund managers most likely possess greater knowledge when it comes to Swedish companies and how the macroeconomic environment is currently affecting the local stock market, and as a result, the portfolios of their mutual funds have a relatively greater weight towards the Swedish market as compared to the rest of the world. The large exposure to the American stock market can be explained by the sheer size of the U.S. stock market, which is the largest in the world. Since most of the largest and most successful companies in the world are traded on the U.S. stock market, a great number of investors, including Swedish ones, are

interested in it. Subsequently, it is not surprising that many Swedish mutual funds provide alternatives where Swedish investors can get exposure to these markets.

Following this line of reasoning, the argument for utilizing Swedish FFC factors, initially described in *Section 3*, is further strengthened since, as mentioned in the paragraph above, there is an over-indexing towards the Swedish market. Thus, the Swedish FFC factors are more relevant and applicable to our mutual fund dataset compared to what an alternative geographical version of the FFC factors would provide. For instance, if FFC factors related to Europe as a whole were utilized, since the exposure to the European market, as seen with *BMF1*, has negative beta values, it would perform poorly in comparison. Naturally, when performing comparisons of performance, they will be more accurate when the market exposure of both the mutual funds that make up our dataset and model factors are the same, which we achieve by utilizing the Swedish FFC factors as well as benchmark factors tuned to such an exposure. It should be noted, however, that the comparatively large exposure to the U.S. stock market could validate the choice of an American set of FFC factors, but as seen in the above-mentioned tables, exposures of several kinds exist, and not all of them can be included in such factors. The choice of Swedish FFC factors, as well as the designated benchmark portfolio, was a compromise based on the fact that we believed that exposure to the Swedish market would be most relevant, and, as can be seen with the betas for *BMF2* and *BMF8*, such an indication is prevalent. In addition, the Swedish economy should be sufficiently interconnected with the American economy, being a major trade partner, to provide representation for such factors at a satisfactory, albeit not perfect, level with regard to the conducted analysis.

Overall, this seems to confirm that the compromises made when choosing the methodology and related datasets were justified, which subsequently strengthens our confidence in the accuracy and trustworthiness of our results and the conclusions drawn from them.

D *Implications*

Following that which has been expressed in the previous parts of the section, it becomes apparent that for the purposes of this paper, the results stemming from the models utilizing solely the alpha measurement, having an overwhelming majority of insignificant results, are inept in their current form, and for this set of data, in evaluating that which was set out in the posed hypotheses. In stark contrast to the alpha models are the value-added models, which, although not perfect in their own right, improve in this aspect to such a degree that it is possible to draw concrete conclusions from them. Before moving further, however, it seems prudent to elaborate on the adjusted R-squared values presented. As was presented in *Section 5* and developed earlier in this section, they are all in the realm of giving a moderate explanatory value to all four models, which is to be expected given the circumstances. Furthermore, it could also serve as an indication that other factors are at play, including the potential for managerial performance. However, it should be noted that managerial performance is unlikely to make up the entirety of the unexplained variance in these models, and it is likely that other unidentified factors make up the majority of it. Nonetheless, this potential for the existence of managerial performance, in combination with the implications from the other results as presented previously, is intriguing.

Expanding on this discussion, in interpreting the results of the value-added models, a couple of things are perceivable. For one, both models exhibit a somewhat

cohesive and discernible pattern stemming from the alpha values and value-added values with respect to the ranking of their respective sets of portfolios. Portfolio 1 performs the best of the ranked portfolios, and Portfolio 10 the worst, a consistent finding for these models and this measurement. This essentially proves that the ranking procedure, as described in the methodology, works as intended, which subsequently implies that the value-added measure is effective in predicting and measuring the performance of the mutual funds. Furthermore, the Long-Short portfolio, an additional indication of such an interpretation, consistently performs the best out of all portfolios in these two models. Essentially, the results indicate that there are unexplained performance gains and losses in the mutual funds within the dataset, both with regard to alpha and value-added. These gains are then consistently predicted and measured by the two models utilizing the value-added measurement, which together with the adjusted R-squared interpretation makes it a reasonable observation that mutual funds on the Swedish market benefit from managerial skills.

The next development is to ascertain which of the two methods of calculating the value-added, namely the benchmark factors or the FFC factors, is most effective. A case can certainly be made for both of them, but as was described when interpreting the statistical measures, only the FFC factors performed with regard to the significance of the results when looking at the lower half of the portfolio set. As it is not possible to solely rely on the portfolios performing well, the importance of Portfolio 10 in particular cannot be understated, and the risk of committing a Type I error when utilizing the results for these portfolios stemming from the benchmark value-added model is deemed too great, it cannot be confidently used in answering our hypotheses. Together with the earlier implications, this then leaves the FFC value-added model as the sole version fitting in accurately predicting and measuring the overperformance and underperformance of the Swedish mutual funds making up our dataset, as well as the subsequent managerial performance, within the context of the performed analysis.

E *Linkages to Previous Literature*

Our paper, as mentioned in the introduction, is inspired by the works of Carhart (1997) and Berk and van Binsbergen (2015), and it could therefore be interesting to compare the results as derived in this paper with theirs, even though they are not directly comparable. To elaborate, since we have not used the same datasets and the methodology, as a result of being a modified combination of the two, is not directly comparable to either, comparing the results directly between the papers will not provide any concrete answers with respect to potential conclusions. Rather, the reasoning behind this comparison is to get an understanding of the reasonability of the general trends or the direction of the results as presented above, albeit at differing magnitudes.

Looking at the monthly excess returns and standard deviations between our models and Carhart's (1997), we find that our values are significantly lower compared with the values reported in *Table 3* from Carhart's paper. The reported monthly excess return for his 1 (high) portfolio is, for instance, 0.68% meanwhile the highest monthly excess return we have for any portfolio out of the four models is significantly lower at 0.1660%. Additionally, the standard deviation Carhart found is anywhere between 4 and 5 percent, meanwhile, our range is roughly between 0.4 and 0.5 percent for all our models. The reason for this difference in both the size of returns and standard deviation, which are inherently related, could be the result of different market conditions between the U.S. and Sweden. In addition, the different time periods of the datasets might have had an

effect due to the possibility of higher returns in earlier years of the data, which could be why Carhart found larger monthly returns.

Furthermore, in observing the adjusted R-square values between Carhart's model and our four models, it is evident that our models have a much lower score compared with his. The range in his paper is approximately between 0.8 and 0.98, while ours is roughly between 0.2 and 0.4. The only common trend between our adjusted R-square values is related to the Long-Short portfolio, where both have a considerably lower value compared with the rest of the portfolios. The fact that our regressions have lower adjusted R-square values is in itself not a problem per se, as it is notably difficult to predict mutual fund returns, thus values between 0.2 and 0.4 should be considered reasonable considering the context of the analysis. Explaining why Carhart's paper reports higher adjusted R-square results overall could be related to how U.S. mutual funds tend, to a greater extent, only be exposed to the U.S. market, while our Swedish mutual funds, as previously described, have exposure to Swedish, U.S., and other markets. Subsequently, they are therefore less homogeneous compared to their American counterparts. Although U.S. funds undoubtedly also have global market exposure to some extent, in general, this is deemed to be smaller. Thus, effectively capturing mutual fund performance using regression analysis is deemed significantly harder when utilizing Swedish factor portfolios such as those utilized in this paper.

Continuing, and looking at the value-added models and comparing them with the results of Berk and van Binsbergen, we find that in our case, around half of our mutual funds produced a positive value-added in the FFC version, which is roughly equivalent to the 41% of mutual funds that produced a positive value-added in Berk and van Binsbergen's case when they used FFC factors. An interesting thing to note, however, is that the difference between positive value-added results between the model based on the benchmark factors and the model based on the FFC factors is only 2% in Berk and van Binsbergen's paper, but in our paper, the difference is 20%. One possible explanation for this can be found in the construction of our benchmark portfolios. Connecting to the discussion in *Subsection 3B* and *Table 2*, our benchmark portfolio exhibits significantly higher fees compared to the Vanguard benchmark utilized by Berk and van Binsbergen. Additionally, there are fewer funds in our benchmark, and some of the exposure to markets that the Vanguard benchmark encompasses is not available in Sweden and therefore absent. Put differently, the large difference between the value-added results stemming from the benchmark factors and the FFC factors in our models could potentially be explained by the inefficiencies of our constructed benchmark.

Lastly, it could be interesting to look outside the direct inspirations and toward the other papers mentioned in the theoretical background. A largely influential paper in the research area of this paper is Fama's (1970) paper on the efficient market hypothesis, and our results could be interpreted as refuting the claim that it is not possible to consistently beat the market by finding mispriced assets. Simultaneously, it should be noted that the above statement of performance should be interpreted as being post any potential fees charged by the mutual funds, as the analysis performed in this paper excludes such a measure of performance, that being the net alpha, a distinct conclusion in this regard is thus difficult to obtain given the circumstances. Nonetheless, it can be loosely stated that the results of this study more closely resemble those not refuting the existence of a potential mutual fund managerial skill than those denying its existence. One such paper previously mentioned is Grinblatt and Titman (1992), which also found evidence indicating the existence of such skill. The results are also, to some extent, in line

with Fama and French (2010), who found indications of managerial performance and subsequent skill when looking at gross measures. It is prudent, however, to note that when looking at net measurements, they did not find such indications, something that this paper has difficulty refuting.

7 Limitations

It is important to note, as related to the discussion in the previous section, that there are some key problems and limitations that might have affected the overall outcome of the paper, its subsequent results, conclusions, and their inherent credibility.

In looking at the amount of available data for Swedish mutual funds, one sees that it is significantly reduced when compared to the datasets used by both Carhart (1997) and Berk and van Binsbergen (2015). Due to several reasons laid out in the data section, our data stretches from January 2006 to December 2019, a period of 14 years, which is significantly shorter as compared to the 30 and 50 years of data that the aforementioned papers had at their disposal. Additionally, a lot of the mutual funds in our datasets are incomplete when looking at the entire period. There are three main reasons for this: the fund either became obsolete before the end of the chosen time frame; the fund was not established at the start of our time frame; or there is a lack of reported values. For that last part, essentially, before data clean-up and processing, in certain periods for certain funds, data was not reported monthly but bi-monthly, or even quarterly. Furthermore, for certain funds, data was simply missing without any discernible pattern or explanation. The reasons behind this phenomenon cannot be confidently explained, but one theory is that reporting standards have evolved in our chosen time frame. Nevertheless, in ensuring data accuracy, we shortened the timespan of our data as much as possible, but we were still left with a few missing values related to both sources, due to which we had to utilize a different method to deal with them, namely filling them out with linear interpolation. This decision was based on several factors and considered trade-offs before being subsequently deemed the best. However, we are well aware of the limitations this may impose on the potential conclusions we can draw from the analysis conducted. In the end, the consequences of both the lesser data availability, the measures taken to fill out missing intermittent values, and other factors mentioned previously are that the reliability of our regressions and the models on which we built our analysis becomes less rigid since there is a lesser amount of data available to run them on, which has been a consistent issue throughout the research process.

Another area where data availability has been an issue is when constructing our eight Swedish fund benchmark factor dataset. As described in the data section, we did not have the available returns for all funds for the entire period specified. Therefore, we had to create an orthogonal basis set as described in the methodology. This follows Berk and van Binsbergen (2015), however, our dataset contained a significantly larger number of missing values than theirs, meaning that the result, namely the benchmark factor dataset is not as statistically credible as theirs, which subsequently affects the analysis by lowering the reliability of the regressions. Furthermore, when creating the benchmark, the fees our Swedish funds charged were significantly higher when compared to the fees charged by Vanguard index funds used by Berk and van Binsbergen. This is because Swedish index funds tend to charge higher fees on an aggregate level as compared to U.S. alternatives, even though both are passively managed. Moreover, some of the funds included in the benchmark factor portfolio are not strictly speaking index funds, but rather

so-called “index-near” funds, which essentially means that they exclude certain stocks based on ESG factors as deemed by the company behind the fund. For a further explanation of what an index-near mutual fund is, we refer back to *Subsection 3B*. The combined implication this can have for our results is that our constructed benchmark might not be perfectly representative of what an investor might choose as a passive alternative investment strategy consisting of entirely passive index funds. Simultaneously, constructing a perfect equivalent to their benchmark was deemed impossible within the context of our research, and as such, our benchmark is deemed to represent what was easily available to Swedish investors during the period. Put differently, our constructed benchmark is a compromise between selecting as cheap, widely traded, and passively managed funds as possible while at the same time adhering to the market exposures as set out by Berk and van Binsbergen. In the end, we risk that some of the returns in our benchmark might not be accurately representative of an entirely passive investment strategy in reality, and therefore it might compromise the credibility and comparison of the results and subsequent conclusions.

Furthermore, as mentioned in earlier parts, there are problems inherent with using a Swedish version of the FFC factors, such as the ones utilized in our paper. A better alternative would therefore be if we could have used a combination of U.S. and Swedish FFC factor values, as those were the largest individual market exposures in our mutual fund dataset. If we had used them, we believe that our results for the FFC models would have been even more accurate. However, such a combination is, to our knowledge, not practically possible, and as such, we deemed the Swedish FFC factors the most fitting compromise.

Much of what has been brought up above can also be interpreted as having led to the reported issues of significance and explanatory values of the models. The importance of shedding light on this cannot be understated, especially considering that even the FFC value-added model, used as a primary piece of evidence in motivating our conclusions, suffers from significance issues with several of the results reported for its respective portfolios. Although these concerns were to some extent laid to rest given the circumstances of the analysis, they are still an issue that plagues it and its reliability, which is something to keep in mind moving forward.

Lastly, it is also important to note that outside of these above-mentioned procedural problems, there is also one significant structural issue that, at the very least, has affected the implications of the conclusions of this analysis, that is worthwhile mentioning, namely the exclusion of the net alpha as a measurement. This is motivated by several factors, including limitations to the scope of the paper, the methodologies used by our inspirations, and others. Nonetheless, excluding the net alpha measure severely limits the implications of the results in practical applications. In other words, if it cannot be determined what part of the abnormal returns earned by mutual funds is actually obtainable by investors, then outside of academia, such results have little meaning. Having said that, since the reported gross alpha values, as seen in *Tables 3, 5, 7, and 9*, are all significantly below the fees charged by our passively managed index funds, as seen in *Table 2*, then one would assume that the fees charged by actively managed funds should be higher and therefore that the amount obtainable by investors is less than zero. However, as this cannot be definitely proven, this paper nevertheless carries mostly theoretical implications.

8 Conclusion

In this paper, we set out to perform a comparative analysis on the subject of mutual fund performance as predicted and measured by the gross alpha and the value-added measurement. In doing so, three distinctive hypotheses were posed surrounding the subject. The previous sections of this paper explore the motivations, conduction, and results of an analysis aimed at proving these hypotheses and expanding the available literature in the area. At last, this paper has arrived at applying what has been discovered to the initially posed hypotheses.

Starting with the first hypothesis, or *H1*, which relates to the existence of mutual fund managerial skill, we conclude that it does exist and thus reject the null hypothesis that it cannot be observed. From the results in *Section 5* and as mentioned in *Section 6*, the value-added models reveal a clear, persistent pattern where the higher-ranked portfolios demonstrate both greater monthly excess returns, larger alpha values, and larger value-added values as compared to their lower-ranked counterparts, providing a strong argument for the existence of managerial skill. Taking into consideration that the adjusted R-squared values of the respective models left a considerable amount of unexplained performance gains and losses of the mutual funds within the dataset, it seems reasonable to assume that at least parts of the unexplained performance can be attributed to managerial skill. This stems from the fact that managerial skill is deemed a major factor affecting the characteristics surrounding a mutual fund as well as its positioning in the market. This conclusion does not change even if this pattern is not visible for the alpha-sorted models, as the results related to those models were deemed insignificant.

Following, the second hypothesis, or *H2*, concerns the relative performance of the alpha measurement versus the value-added measurement in evaluating and predicting mutual fund manager performance. As has been derived in the previous sections of this paper, especially the discussion in *Section 6*, the models related to the value-added measurement performed significantly better in all aspects as compared to those derived from the alpha measurement. Moreover, the results of the models related to the alpha were deemed as largely insignificant and those stemming from the models related to the value-added figure were deemed sufficiently significant for the purposes of this analysis. Thus, given the confirmation of *H1* above, it seems reasonable to reject the null hypothesis as related to *H2* and declare, for the purposes of this paper and the analysis performed within it, that the value-added measurement is the better predictor and measurement of Swedish mutual fund manager performance.

Lastly, the third hypothesis, or *H3*, has to do with whether or not the benchmark portfolio factors are the better predictor as compared to the FFC factors. Our benchmark value-added model, as described in *Section 6*, has problems when it comes to the significance levels for the lower-ranked portfolios, and thus, we cannot be certain of the accuracy of the results stemming from it. The reason for this poor significance level could be attributed to the issues related to the construction of the benchmark, and thus the benefits of using the benchmark as advocated by Berk and van Binsbergen in their paper are not evidently proven in our case. In contrast, the FFC value-added model performs slightly better, namely because both the highest and lowest portfolios, which are the most important portfolios, are significant. The portfolios in the middle are, however, statistically insignificant. It is thus not possible to determine which of the two models is the better predictor, and therefore we fail to reject the null hypothesis, meaning that, for

the purposes of this paper, the FFC factors, serving as the default option, are the better predictors of Swedish fund managerial performance.

In conclusion, the first, *H1*, and second, *H2*, hypotheses could be accepted, while the third, *H3*, could not and was subsequently rejected. That is to say, the results stemming from the performed analysis have, within its context, confirmed the existence of Swedish mutual fund managerial skill as well as accepted the value-added measurement as the superior predictor and metric in evaluating it. However, the results were insufficient to generate any concrete conclusion to answer the question of whether the benchmark portfolio factors were better than the FFC factors for that same purpose. Thus, the default option, as well as that which was related to the stated null hypothesis, namely the FFC factors, are deemed superior for the purposes of this paper. These findings are to some extent comparable to those achieved by Berk and van Binsbergen (2015), who also confirmed the existence of fund managerial skill, as well as deemed the superior measure in doing so to be the value-added measurement when looking at the U.S. mutual fund market. It should be noted that the acceptance of *H1* and *H2* ought to be considered conditional in light of the problems plaguing the conducted analysis, as mentioned in *Section 7*. Most of these issues stem from the structure and availability of data related to the Swedish mutual fund market; in particular, the prevalence of missing data points has potentially affected the quality of our regressions. Despite efforts made to clean and transform the data, there are still issues left that could not be avoided. Furthermore, the much smaller offering of financial products, particularly when looking at products tracking diverse indexes, has also had a detrimental effect on the analysis, as the construction of the benchmark, a core aspect of the analysis, was severely limited. Moreover, it is also the case that the conducted analysis omits the net alpha measurement, leading to any practical implications being noticeably limited in scope. Hence, we believe that further exploration in this field of Swedish mutual fund manager performance would be appropriate to strengthen the findings of this paper; however, as has become evident in this paper, the data availability might make such a venture difficult, which is why we propose looking at additional data sources or venturing to other closely related markets, such as the other Nordic countries in order to draw more far-reaching conclusions. Building on these suggested extensions and expanding the analysis to include the presently absent net alpha measurement, the practical implications would also be greatly improved. In doing the above, albeit not perfectly, we believe that the analysis would most optimally be adapted and expanded to move past the identified issues as well as make further advances in this area of research.

9 References

- Barber, B. M., Huang, X., & Odean, T. (2016). Which Factors Matter to Investors? Evidence from Mutual Fund Flows. *The Review of Financial Studies*, 29(10), 2600-2642.
- Berk, J. B., & van Binsbergen, J. H. (2015). Measuring skill in the mutual fund industry. *Journal of Financial Economics*, 118(1), 1-20.
- Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, 52(1), 57-82.
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417.
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56.
- Fama, E. F., & French, K. R. (2010). Luck versus Skill in the Cross-Section of Mutual Fund Returns. *The Journal of Finance*, 65(5), 1915-1947.
- Grinblatt, M., & Titman, S. (1989). Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings. *The Journal of Business*, 62(3), 393-416.
- Grinblatt, M., & Titman, S. (1992). The Persistence of Mutual Fund Performance. *The Journal of Finance*, 47(5), 1977-1984.
- Grinblatt, M., & Titman, S. (1993). Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns. *The Journal of Business*, 66(1), 47-68.
- Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, 48(1), 65-91.
- Jensen, M. C. (1968). THE PERFORMANCE OF MUTUAL FUNDS IN THE PERIOD 1945-1964. *The Journal of Finance*, 23(2), 389-416.
- Sodini, Paolo, Yao Fu, and Huseyin Aytug. (2020). Construction of the Fama-French-Carhart Four Factors Model for the Swedish Stock Market Using the Finbas Data. Swedish House of Finance Research Data Center. Retrieved February 10, 2024, from <https://www.hhs.se/en/houseoffinance/data-center/fama-french-factors/>
- Statistiska Centralbyrån. (2024, March 15). CPI, historical numbers, 1830-. Retrieved March 15, 2024, from <https://www.scb.se/en/finding-statistics/statistics-by-subject-area/prices-and-consumption/consumer-price-index/consumer-price-index-cpi/pong/tables-and-graphs/consumer-price-index-cpi/cpi-historical-numbers-1830/>
- Wermers, R. (2000). Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses. *The Journal of Finance*, 55(4), 1655-1695.

10 Appendix

Disclaimer Surrounding the Usage of Artificial Intelligence: In the making of this paper, the large language models (LLMs) GPT-3.5 and GPT-4 by OpenAI were utilized as assistance during the making of code used in the analytical regressions. Furthermore, such models were also used in the context of verifying the function and results of such code to ensure the stability of our analysis. In addition, the AI-enabled writing assistant Grammarly was continuously used to check for grammatical errors during the writing process.