ERASMUS UNIVERSITY ROTTERDAM

Erasmus School of Economics

Master Thesis Financial Economics

What factors do investors care about? Evidence from the European mutual fund market.

Name Student: Gino Laurentius Antonius Pel

Student ID: 452367

Supervisor: Ricardo Barahona

Second assessor: Matthijs Korevaar

Date final version: July 27, 2021

Acknowledgements

As this thesis marks the end of my time as a student, I would like to thank first and foremost the love of my live for supporting me throughout the final phase of my student life and for supporting me in all my future endeavors. I would like to thank my parents, sister, and grandparents for their everlasting support throughout my life, without that I would not have come so far. I also want to thank Harry and Smebby, for always being present and showing me the simple joy of life.

Finally, I of course want to thank my supervisor Ricardo Barahona for answering all my questions and providing guidance in the thesis process.

Abstract

A recent surge in the mutual fund flow literature investigates what factors matter to investors when assessing a fund managers performance. The U.S. evidence on this topic is mixed, with studies suggesting that investors use a basic Capital Asset Pricing Model, while another study finds these results to be spurious and instead shows that the evidence is most consistent with investors blindly following Morningstar Ratings. This thesis sheds light on this matter by investigating the European mutual fund market, and finds that investors account most for market risk, whilst treating size-, value and momentum-related return components as alpha. It is tested whether these results are spurious, which does not appear to be the case for the European mutual fund market. However, the results also show that European mutual fund investors primarily let their investment decision be influenced by Morningstar ratings. In general, European mutual fund investors use less sophisticated asset pricing models in their investment allocation decision and appear to be less than fully rational by using attention grabbing ratings.

Table of Contents

1. What factors do investors care about? Evidence from the European mutual fund market.

What factors do mutual fund investors care about when allocating their savings to actively managed mutual funds? Normatively, a rational investor should use all available information and consider all factors that explain cross-sectional variation in fund performance to determine which manager is skilled. A recent surge in the literature provides mixed evidence of how investors descriptively behave. On one hand, U.S. mutual fund investors appear to use a basic Capital Asset Pricing Model (CAPM), while other studies find that investors behavior is most consistent with blindly following Morningstar ratings and chasing past returns (Barber et al., 2016; Ben-David et al., 2019; Sirri & Tufano, 1998). These studies have been limited to the U.S. market, while at the same time the European mutual fund market remains under researched in the finance literature (Banegas et al., 2013). This study aims to shed light on this question and fill this gap by investigating what factors matter to European mutual fund investors when assessing a fund manager's performance.

The relevance for filling this literary gap of the largely unexplored European mutual fund market is underlined by its growing importance over the past decades. The European mutual fund industry as of 2020 amounts to a total net asset value of 21.8 trillion USD, 35% of the global industry. Moreover, the empirical research done in the U.S. need not apply to the European mutual fund market, as both markets have different fundamental characteristics. For instance, in the U.S. nearly half of households (47.4%) invest in mutual funds, and they hold a two and a half times larger portion of their wealth in regulated funds than the more bank-centric European market. The regulated fund centric U.S. is in part due to the mutual fund industry originating from the U.S, having existed for almost 100 years. Additionally, long standing defined contribution plans, such as the 401-k, has influenced many households to invest themselves (ICI, 2021). Besides this, the past studies had data until 2011. By studying European mutual funds, in addition to adding the recent 10 years of data, this study can provide a novel insight into the investment decisions of investors.

A recent strand of the finance literature studies mutual fund flows, the elegance of mutual fund flows is that it reveals investors preferences. Fund flows are the aggregate cash in- and outflows into mutual funds, representing buy- and sell order from investors. With that in mind, it should be considered that the vast majority of mutual fund assets is held by retail investors[1]. In a seminal study, Barber et al. (2016) investigate whether mutual fund investors tend to common factors and industry tilts by measuring the sensitivity of flows to different factor-related fund return components. They find that investors tend most to the market risk factor, treating other factor-related returns as alpha. To preview the results, the analysis in this study finds similar results for European mutual fund investors with one important distinction; European mutual fund investors appear to fully account for the market risk factor. The authors interpret this as investors having different levels of sophistications, with retail investors

[1] In the U.S., 89% of mutual funds total net assets is held by retail investors as of 2020 (ICI, 2021).

being less sophisticated. A sophisticated investor should attend to all factors that explain cross-sectional variation in fund returns. Viewed from this lens, this study assesses the financial sophistication of European mutual fund investors by investigating what common factors matter to those investors.

A later study by Ben-David et al. (2019) re-examines the results of Barber et al. (2019) and argue that those results lead to spurious evidence in favour of investors using the CAPM. The authors argue this using a new empirical fact, first documented by Franzoni and Schmalz (2017) in the U.S., that the flow-performance sensitivity (FPS) varies over time. Ben-David et al. (2019) instead find that fund flows are most consistent with investors blindly following Morningstar ratings and chasing past returns. This inference has support from the literature documenting that investors are less than fully rational, and that retail investors are particularly susceptible to biases and attention-grabbing signals (Del Guercio & Tkac, 2008; Sirri & Tufano, 1998; Solomon et al., 2014). Although Morningstar is less ubiquitous in Europe than in the U.S., the results found by Ben-David et al. (2019) could have a profound impact to the research question at hand. This study will, therefore, answer whether the investment decisions of European mutual fund investors are influenced by attention grabbing signals such as Morningstar ratings.

To test whether these results might be spurious, as argued by Ben-David et al. (2019), I first document that the time-varying FPS in the European mutual fund market is much less pronounced than in the U.S. market. Then, a formal test is performed using simulated flows which by construction do not differentiate between fund-return components. The results of the tests generally show that the coefficient estimates in the main panel regression are not found mechanically and that the general inference still hold. The results are also not explained by Morningstar ratings, meaning that within each star rating investors still mainly account for the market-related fund return component. To more rigorously test the ability of Morningstar ratings to explain the capital allocation decision of investors, the flows to top- and bottom ranked funds of different asset pricing models and star ratings are analyzed. Remarkably, the Morningstar ratings far outperform any other performance measure in predicting flows. Top Morningstar rated funds receive an average of 3.6 billion USD more flows annually than the next best performance measure. Taken together, it can be said, with a greater degree of certainty than can be said in the U.S. market, that European mutual fund investors account for market risk, whilst at the same time, largely let their investment decision be influenced by Morningstar ratings.

The rest of this paper is organized as follows. Section 2 explains the theoretical background of this research. Section 3 explains the data collection process and the methodologies used for the main panel regression analysis, after which the data is described. Section 4 presents the results, in which first the flows to top-and bottom ranked funds is analyzed. Second, the results of the main panel regression of flows to factor related return components is presented. Third, it is tested whether the time-varying FPS holds in the European market and whether the main panel regression results are biased. Fourth, a robustness test on the decay rate parameter of the independent variable in the main panel regression is performed. Finally, section 5 concludes this study.

2. Theoretical background

This research fits within the mountainous body of literature on mutual funds. This area gets a lot of attention in the financial literature because of its sheer size and impact on the economy. Moreover, it provides ample high-frequency data, with which researchers are equipped to answer a multitude of questions. Some of the early literature investigates the performance and persistence of fund managers. Jensen (1968) is one of the first to document the general underperformance of mutual funds. Studies since then have provided ample evidence showing that funds on average underperform their benchmark and that there is little persistence in that performance over time (Carhart, 1997; Fama & French, 2010; Malkiel, 1995). A more recent strand of literature studies the money flows to mutual funds. The elegance of studying mutual fund flow data is that it reveals investors preferences. Earlier works in this area document that the flow-return relationship is generally convex, positive returns receive more inflows than the same negative returns receive in outflows, the so called return-chasing behavior of investors (Chevalier & Ellison, 1997; Sirri & Tufano, 1998). Sirri and Tufano (1998) argue that search costs are an important determinant for fund flows. As the attention effect entails, investors have a limited amount of attention they can devote to investing. Investors generally do not sift through the thousands of fund options available to determine which investment is best, but instead buy attention-grabbing funds, like the funds with extraordinary past performance. Later studies further investigate investors preferences by studying the relationship between flows and risk. Clifford et al. (2013) provide evidence that both fund inflows and outflows are positively related with a fund's total risk. They find that primarily for retail investors, the inflows are positively related with idiosyncratic risk. The authors argue that retail investors chase past idiosyncratic risk as they naively extrapolate past returns into the future. This adds evidence to the earlier studies, documenting the return-chasing behavior of investors, that mutual fund investors investors tend to behave sub-optimally.

Going beyond the relation between flows and total returns or total risk, two fundamental recent studies investigated the relationship between flows and commonly used factors that explain cross-sectional variation in performance. Some of these factors are generally acknowledged to be associated with risk, such as the market risk factors. However, there is still controversy about which of the other common factors constitute as risk (take the value factor for instance). Therefore, this study abstains from calling all of these factors 'risk factors'. These recent studies, and the discussion that followed in the literature, are the main source of inspiration for this research. Both Barber et al. (2016) and Berk and Van Binsbergen (2015) independently reach the same conclusion that fund flows are best explained by investors using the Capital Asset Pricing Model (CAPM) model to assess fund managers performance. In other words, investors appear to understand that funds have a certain exposure to the market risk factor, and that fund returns should be adjusted for the returns attributable to the market risk factor when determining a fund managers skill. Both studies interpret these results in a mutually exclusive manner. Barber et al. (2016) argue that these results indicate that retail investors are generally unsophisticated, as they do not consider all factors when assessing a fund managers skill. The unsophisticated investor confuses managers skill with returns attainable through exposure to common factors using passive investment. Berk and Van Binsbergen (2015) argue that their empirical analysis is a test which can reveal the 'true' asset pricing model. They interpret their results as the CAPM being the closest to the true asset pricing model for *all* investors and that it is the best asset pricing model practitioners should use. This bold inference seems intuitively dubious. Barber et al. (2016) argue that this inference cannot be made for non-mutual fund investors and Jegadeesh and Mangipudi (2021) go even further to show that a faulty foundational assumption leads them to make this mistaken inference. Overall, I agree with most other researchers and regard the interpretation of Berk and Van Binsbergen (2015) to be wildly inappropriate. Consistent with the motivation of this study, I regard these empirical analyses as revealing what factors European mutual fund investors care about when they allocate their funds.

Both Barber et al. (2016) and Berk and Van Binsbergen (2015) use an empirical flow-alpha horse race to test which of the competing asset pricing models does the best job at explaining variation in flows across funds. Ben-David et al. (2019) argue that the results found by Jegadeesh and Mangipudi (2021) challenge the validity of the empirical flow-alpha horse race test, as it favors an asset model with fewer factors such as the CAPM. This argument from Ben-David et al. (2019) seems to be self-serving and not correct, however, as Jegadeesh and Mangipudi (2021) clearly state that even with the four-factor model as the benchmark, the result that the CAPM alpha wins the empirical race suggest the rejection of the investor sophistication hypothesis[2]. An actual downside of this test is that it can only tell which asset pricing model does the best job at explaining fund flows, it is unable to tell 'how much' investors care about factors. In the case of the CAPM winning the horse race, researchers cannot learn whether investors fully account for market risk and completely ignore other factors. To dissect what factors investors tend to, Barber et al. (2016) propose a second test. In this test they investigate how fund flows respond to different factor-related return components. The authors perform this test by first decomposing a funds return into its seven-factor alpha and factor-related return components (market, size, value, momentum and three industry factors), after which fund flows are regressed on those return

[2] Jegadeesh and Mangipudi (2021) explain that from the empiricist view, the true- asset pricing model and betas are unknown, meaning that the model with the most precise alpha estimated *always* wins the empirical horse race under the rational expectations hypothesis. According to the authors, the model with the most precise alpha estimate and thus the winner of the horse race in this context is the four-factor model. What Jegadeesh and Mangipudi (2021) *do* dispute is the argument of Barber et al. (2016) that a seven-factor model would win the horse race under the rational expectations hypothesis.

components. In this test, they find that investor tend most to the market risk factor, as the flows are least sensitive to this return component. Flows are generally sensitive to all other factor-related return components, meaning investors reward them as if it is alpha. Agarwal et al. (2018) perform the same test for hedge funds to investigate the relation between hedge fund flows and return components. Hedge fund investors, like mutual fund investors, tend most to the market risk factor. Factor-related returns also receive fund flows, although the sensitivity is generally weaker than those found for mutual fund investors. This may indicate that more sophisticated investors use more sophisticated models to assess fund managers skill, as hedge fund investors are generally more sophisticated than mutual fund investors. As this study aims to investigate how European mutual fund investors determine the performance of fund managers, a similar flow-return component analysis will be performed. Giving the prior research, it is predicted that European mutual fund investors tend most to the market factor, whilst generally treating other factor-related return components as alpha.

Ben-David et al. (2019) re-examine the results of the second test performed by Barber et al. (2016) and find that they lead to spurious evidence in favor of investors using the CAPM. Ben-David et al. (2019) instead find that the fund flow data is most consistent with investors following Morningstar ratings and chasing past return. The argument for why the flow-return composition results are likely spurious goes as follows. The flow-return composition panel regression uses time-fixed effects, which overweight's cross-sections in periods of extreme market returns (high or low), as the dispersion in the independent variable, market-related return component, is particularly large. This is key when combined with the time-varying flow-performance sensitivity first document by Franzoni and Schmalz (2017) in the U.S. The authors show that the flow-performance sensitivity (FPS) is a hump-shaped function of aggregate market risk. In times of moderate market returns the FPS is twice or even thrice as large as in extreme market return. As a result, specifically in those extreme market return periods, which largely drive the regression coefficient estimate, the sensitivity of flows to returns is weak. This can, therefore, bias the coefficient estimates downward. Ben-David et al. (2019) show in a formal test that even if investors do not differentiate between factor-related return components, the panel regression would still create results that suggests that flows respond weakly to the market-related return component. This concern raised by Ben-David et al. (2019) could have a substantial impact on the answer to the research question at hand. Therefore, these concerns shall be addressed. The time-varying FPS is not yet tested nor documented for the European mutual fund market. This research will, therefore, investigate whether the European mutual fund market exhibits a time-varying FPS and, more importantly, whether those periods with weak FPS exhibit a large dispersion in the market-related return component. Franzoni and Schmalz (2017) propose a simple theoretical model that best explains the market-state dependent FPS; When systemic risk is muted, performance is more revealing about fund managers skill, causing the increased sensitivity during moderate periods of market returns. Since this explanation should also apply to the European mutual fund market, it is expected that this market also exhibits a time-varying FPS. Additionally, the formal test of Ben-David et al. (2019), which involves simulating flows, is performed for the European mutual fund market.

The finding that fund flows are most consistent with investors chasing Morningstar rating, has extensive support from the literature which documents that the average mutual fund investors is boundedly rational or even irrational. Sensoy (2009) document that one-third of actively managed mutual funds has a benchmark listed in the prospectus that does not match the actual fund style. Mutual fund flows react strongly to a fund's outperformance with respect to these mismatched benchmarks. Fund managers respond to these incentives by strategically setting a benchmark. Elton et al. (2002) show that there are investors which invest in high cost S&P500 Index funds, while plenty cheap alternatives exist which can predictably deliver better performance. Choi and Robertson (2020) perform a survey on U.S. households and find that individuals believe that the past performance of mutual funds is a good indicator of stock-picking skill of the fund manager. Moreover, there are studies that provide evidence that investors respond to salient, attention grabbing signals. Frazzini and Lamont (2008) provide evidence that mutual fund investors invest in high sentiment stocks, which predict low returns. In other words, a portfolio that is the exact opposite of individual investors has predictably high returns. Solomon et al. (2014) show that mutual funds with media attention exaggerate the return-chasing behaviour of investors. Most relevant for this research, Del Guercio and Tkac (2008) show that investors respond strongly to Morningstar ratings and rating changes. Motivated by the results of Ben-David et al. (2019) and the expanding literature on the irrationality of investors, it is important and relevant to investigate whether Morningstar ratings are a tool by which European mutual fund investors assess fund managers skill. As the European mutual fund market is less developed than the U.S. market, the Morningstar ratings are likely less relevant in the European market.

Finally, as it is a key element of this paper, it is worth describing the Morningstar ratings so that it helps the ease of discussion. Morningstar is a leading financial services and software company in the mutual fund industry, which according to Del Grucio and Tkac (2008) is the "undisputed leader among retail investors" (p. 908). Its ubiquitous star rating system was first introduced in the U.S. in 1985. As described earlier, the U.S. mutual fund market is further developed and had a head start relative to the European mutual fund market. This is highlighted in the fact that the star rating system was introduced to Europe in March 2001. The rating system is easy to access and process, whilst having a rigorous methodology which updates ratings monthly. The star-ratings, going from 1 through 5 for, respectively, low- and high rated funds, are based on the relative volatility-adjusted performance of a fund within its peer group. A general principle of this relative rating, according to Morningstar, is that the relative rating should reflect a fund managers skill. However, it should be emphasized that the performance measure only adjusts for volatility and does not account for any (risk) factors. These peer groups are defined as one of the Morningstar Categories which groups mutual funds based on their holdings. These, for instance, include one of the nine Morningstar style boxes, which is a combination of size (small, mid, large) and investment style (value, blend, growth), but also include 53 other categories such as Indian Equity. Within each group, 10%, 22.5%, 35%, 22.5% and 10% gain a 1-, 2-, 3-, 4- and 5-star rating, respectively. Before October 2006, only Morningstar ratings over a three-year period (=overall rating) are reported in Europe, after which the three-, five- and ten- year period ratings are reported if the data is available. This research uses the overall rating, which post October 2006 is based on a weighted average rating over the available time-period ratings.[3]

3. Data & Methodology

Data

Data is gathered from the Morningstar Direct database. The survivorship-bias free database has a comprehensive international set of open-ended mutual funds data, which also covers Europe. The key advantage to this database is that it offers Morningstar ratings, in addition to monthly fund returns, total net assets (TNA) and other fund characteristics of interest to this research. To be considered for the sample. a fund must be an open-ended fund, domiciled in Europe, and have the broad category of equity, with at least 75% of total net assets in equity. Index funds are excluded as this study focuses on how investors choose actively managed funds. The dataset contains monthly data beginning in November 1990. This date is chosen as it is the earliest date for which factor portfolio returns are available at Kenneth French's Data Library.

Following the standard methodology in the literature (Barber et al., 2016; Ben-David et al., 2019; Franzoni & Schmalz, 2017), fund flows for fund *p* in month *t* is defined as the percentage growth of lagged total net assets under management, computed as follows

$$F_{p,t} = \frac{TNA_{p,t}}{TNA_{p,t-1}} - \left(1 + R_{p,t}\right) \tag{1}$$

Here, $TNA_{p,t}$ is the total net assets under management of fund *p* in month *t*, and $R_{p,t}$ is the return of fund *p* in month *t*. Flows that exceed -90% or 1,000% in each month *t* are trimmed from the data. Fund's individual share classes are then aggregated to the fund portfolio level. Returns, expense ratio and fund flows (both in percentages and in dollars) are computed as the share class TNA weighted average across share classes. The same is done for Morningstar ratings, which are subsequently rounded to the nearest integer. For the other control variables, a dummy for no-load is true if all share-classes are no-load funds, meaning they have no front- or back-end load. Fund age is computed as the time between the first and last return observation of a fund in the data and volatility is computed as the standard deviation of returns over the given period. Volatility and expense ratio are winsorized at the 1st and 99th percentile to remove extreme outliers.

[3]https://www.morningstar.com/content/dam/marketing/shared/research/methodology/771945_Morningstar_Rating_for_Funds_Methodology.pdf

Sample inclusion requires TNA and expense ratio data at month *t*, at least 10 million USD TNA at month *t* – 1 and non-missing Morningstar rating data as to maintain comparability across analyses. The analysis is also limited to funds which have month *t* – 60 to *t* – 1 non-missing return data, which is required in order to estimate factor return components. This criterion has the added benefit that the first five years of a fund's existence are excluded from the sample. The consequence is that the sample is free from incubation bias, which otherwise could affect inference (Evans, 2010)[4]. The main sample consists of 3,621 unique European open-ended mutual funds and covers nearly 20 years, beginning in September 2001 and ending in April 2021[5].

Mutual fund performance and return decomposition

The four-factor model (Carhart, 1997) is used to infer how European mutual fund investors respond to different components of fund returns. The four-factor model is chosen over the seven-factor model which adds three industry factors, used by Barber et al. (2016), since four-factor alphas are more precise than seven-factor alphas when true betas are unknown (Jegadeesh & Mangipudi, 2021). This is because there is a large estimation error in the betas of those industry factors. In other words, the four-factor model is more precise in estimating the true skill of a fund manager. For those interested, Barber et al. (2016) find evidence that investors do not attend to the three industry factors. The four-factor model is also chosen over promising new models, such as the Q-factor model (Hou et al., 2017), for the simple fact that this model was first introduced in 2015, while the sample starts in 2001. Meaning that even sophisticated investors would likely not respond to some of the factors of the Q model, since they would not have been able to use it to judge fund manager performance. Four-Factor return data is obtained from Kenneth French's Data Library[6]. This is source is well-regarded in the literature and has the additional benefit of being freely available and easily accessible, meaning that investors may actually use this data to determine their investment decisions. Since the European domiciled funds in the sample invest in global and diverse stocks, the factor return data for developed markets is used. For reference, the returns of the factor portfolios during the sample period are presented in Appendix A.

To decompose a funds return, I proceed in two steps. First, a rolling window Carhart (1997) four-factor regression is performed for each fund to estimate factor loadings for each fund-month observation. Then, the factor-related return components are estimated as the factor loading times the factor portfolio return for that month. Alpha is then estimated as the realized return of a fund less the factor-related returns. Formally, for each fund *p* in month *t*, the following time-series regression is run from month *t* – 60 to *t* – 1

[4] Mutual fund families often open multiple new funds with limited capital. After an evaluation period, some are closed down never to be available to investors, while the others become open to the public.
[5] The starting date is due to a combination the requirements of a five-year estimation window and Morningstar ratings at month t, as the first occurrence of these ratings in the data is in 2001.
[6] http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

$$R_{p,t} - Rf_t = a_{p,t} + b_{p,t}(Rm_t - Rf_t) + s_{p,t}SMB_t + h_{p,t}HML_t + w_{p,t}WML_t + e_{p,t} \qquad (2)$$

Where $R_{p,t}$ is a funds return in month *t*, Rf_T is the risk-free rate in month *t* and Rm_t is the return on the market portfolio in month *t.* For month *t,* SMB_t is the return of the size factor portfolio (small minus big), HML_t is the return on the value factor portfolio (high minus low book-to-market), WML_t is the return on the momentum factor portfolio (winners minus losers). The coefficients $b_{p,t}$, $s_{p,t}$, $h_{p,t}$ and $w_{p,t}$ are the market-, size-, value- and momentum factor tilts of fund *p*, respectively. The intercept $a_{p,t}$ is the average return unrelated to factor exposures over the period *t* – 60 to *t* – 1. To estimate the four-factor alpha of fund p in month *t*, I follow the methodology of Barber et al. (2016) and estimate it as the realized return less factor related returns

$$\hat{a}_{p,t} = (R_{p,t} - Rf_t) - \left[\, \hat{b}_{p,t}(Rm_t - Rf_t) + \hat{s}_{p,t}SMB_t + \hat{h}_{p,t}HML_t + \widehat{w}_{p,t}WML_t \right] \qquad (3)$$

Where $\hat{a}_{pt}$ is the four-factor alpha of fund p at month *t*, and $\hat{b}_{p,t}$, $\hat{s}_{p,t}$, $\hat{h}_{p,t}$ and $\widehat{w}_{p,t}$ are the estimated coefficients of Equation (2). The CAPM alpha and the Fama and French (1993) three-factor alpha, used in the analysis of top-bottom ranked performance, are estimated in a similar manner. Rearranging Equation (3) shows how the return of fund p in month t is decomposed in its alpha and factor-related return

$$(R_{p,t} - Rf_t) = \hat{a}_{p,t} + \left[\, \hat{b}_{p,t}(Rm_t - Rf_t) + \hat{s}_{p,t}SMB_t + \hat{h}_{p,t}HML_t + \widehat{w}_{p,t}WML_t \right] \qquad (4)$$

Where, for instance, $\hat{b}_{p,t}(Rm_t - Rf_t)$ is the return of fund p in month t related to the market factor.

Decay function and model specification

A rational investor would update his investments with new relevant information, shifting capital away from poor performing managers and towards the better performing ones. What is not obvious is how an investor weighs past information. Recent returns are likely more informative, while at the same time contain more noise. Longer term returns are likely less informative but are less noisy, creating a trade-off. At the same time, investors may respond slowly to new information (Coval & Stafford, 2007), and as mentioned earlier, at least a sizable portion of mutual fund investors behave in a less than fully rational manner (see, e.g., Barberis and Thaler, 2005; Sirri and Tufana, 1998; Solomon et al., 2014). This complicates the empiricists task of creating an appropriate weighting scheme. The question remains what horizon should be used to evaluate fund performance, and how the past performance should be weighted.

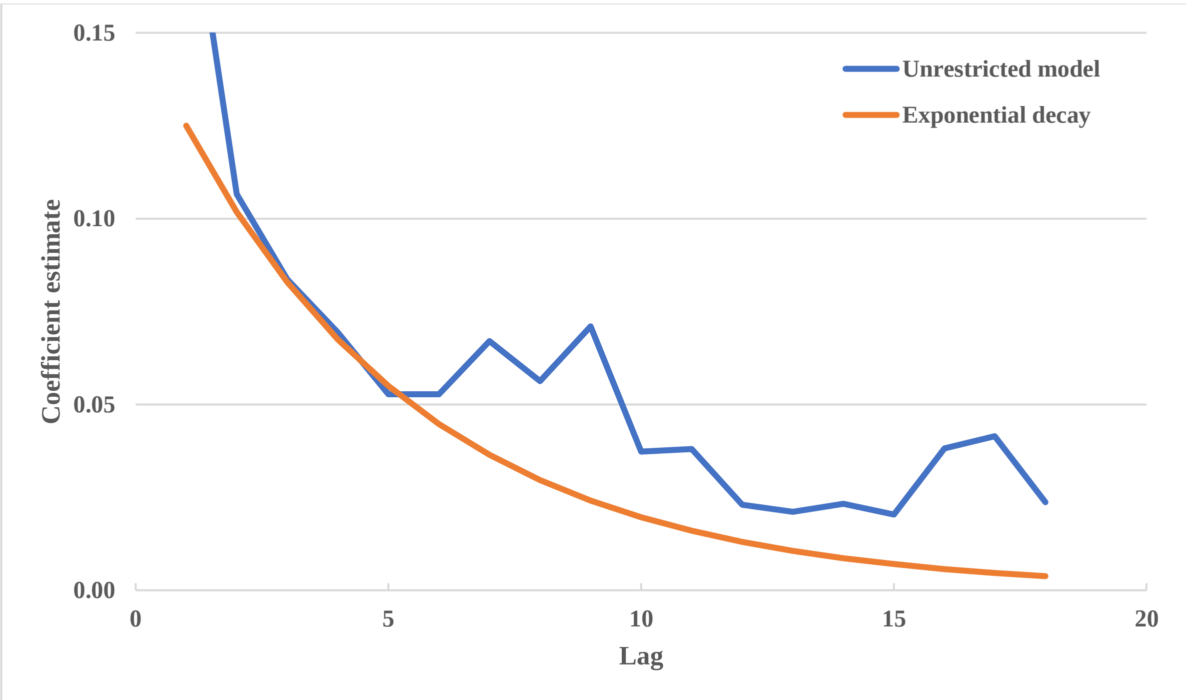
The solution proposed by Barber et al. (2016) is to let the data provide the answer by estimating the empirical rate of decay in the flow-return relationship. By doing so, the response of investors to past performance do not have to be arbitrarily assumed. A potential concern, however, might be that investors of European funds respond differently to past returns than investors of U.S. funds do. To address this concern, I empirically estimate the decay rate in the European mutual fund flow-return relation by estimating the following unrestricted model

$$F_{p,t} = a + \sum_{s=1}^{18} b_s MAR_{p,t-s} + cX_{p,t} + \mu_t + e_{p,t} \tag{5}$$

Here, $F_{p,t}$ are flows of fund p in month t, $MAR_{p,t-s}$ are market adjusted returns of fund p in lagged months $t - s$ where $s = 1$ to 18 months. For comparability, a lag period of 18 months is used, which was the best lag length model based on the Akaike information criterion in the data of Barber et al. (2016). A vector of control variables is captured in $X_{p,t}$ and the vector of coefficients of those variables are captured in c. The control variables include lagged flows for month $t - 19$, a dummy for no-load funds, return volatility over the prior 12 months, and lagged month $t - 1$ values of net expense ratio, log of fund size and log of fund age. In addition, time fixed effects are added, captured by μ_t. The coefficients b_s are the variables of interest and capture the relation between flows in month t and market adjusted returns lagged at $s = 1$ to 18 months.

Figure 1

The decay in the fund-flow relation



This figure presents the panel regression coefficient estimates of the unrestricted model with the dependent variable fund flows on the independent variable of 18 months of lagged returns. This plot shows how sensitive European mutual fund flows to each lag of fund return. The empirically found decay rate $\hat{\lambda} = 0.20551497$ of Barber et al. (2016) by parsimoniously modelling Equation (5), is also displayed for comparison. The regression also includes a set of control variables and time fixed effect. The controls include lagged flows for month $t - 19$, a dummy for no-load funds, return volatility over the prior 12 months, and lagged month $t - 1$ values of net expense ratio, log of fund size and log of fund age.

Plotting the coefficients in Figure 1 shows a clear decay in the European mutual fund flow and past return relationship, where flows are considerably more responsive to recent returns than they are to distant return. This decay rate is near identical to the decay rate found by Barber et al. (2016), with the European fund flows only responding slightly stronger to each lag of returns relative to the U.S. funds. Barber et al. (2016) then proceeds to estimate the flow-return relationship parsimoniously to estimate decay rate λ as follows

$$F_{p,t} = a + b\sum_{s=1}^{18} e^{-\lambda(s-1)} MAR_{t-s} + cX_{p,t} + e_{p,t} \tag{5}$$

Since the fund-flow relation in this context for European mutual fund data closely resembles those found in the U.S. data, there is no concern that European investors respond meaningfully different to past returns than U.S. investors. This study, therefore, does not parsimoniously estimate the decay rate parameter, and instead proceeds with the decay parameter $\hat{\lambda}$ = 0.20551497 found and used by Barber et al. (2016). In a later robustness test, it is shown that reasonable changes in decay parameter λ do not materially affect the main results found in this research.

For the second step, for each fund-month observation, the empirically found decay parameter $\hat{\lambda}$ is used to weight alphas and factor-related returns. For instance, when considering a fund's alpha in month *t*, it is estimated as the prior 18-month weighted alpha in the following manner

$$ALPHA_{p,t} = \frac{\sum_{s=1}^{18} e^{-\hat{\lambda}(s-1)} \hat{a}_{p,t-s}}{\sum_{s=1}^{18} e^{-\hat{\lambda}(s-1)}} \tag{6}$$

Where $\hat{a}_{p,t-s}$ is the estimated four-factor alpha from Equation (3). Each of the factor-related return components are also weighted using this decay function. For clarity, in each fund-month observation, the fund return component related to the market factor is weighted as follows

$$MKTRET_{p,t} = \frac{\sum_{s=1}^{18} e^{-\hat{\lambda}(s-1)} \left[\hat{b}_{p,t-s}(Rm_{t-s} - Rf_{t-s})\right]}{\sum_{s=1}^{18} e^{-\hat{\lambda}(s-1)}} \tag{7}$$

The size, value and momentum factor return components are calculated similarly and are labelled as SIZRET, VALRET and MOMRET, respectively.

With a time-series of empirically weighted return decomposition acquired, it can be inferred how investors respond to different fund return components by estimating the following panel regression

$$F_{p,t} = b_0 + b_1 ALPHA_{p,t} + b_2 MKTRET_{p,t} + b_3 SIZRET_{p,t} + b_4 VALRET_{p,t} + b_5 MOMRET_{p,t}$$

$$+ \gamma X_{p,t} + \mu_t + e_{p,t} \tag{8}$$

Where $F_{p,t}$ are flows to fund p in month t. Three different models are used, where the only parameter that changes is μ_t. In the first model μ_t is month fixed effects, in the second model μ_t is month-style (two-way) fixed effects and in the third model μ_t is month-style-rating (three-way) fixed effect. Here, rating is the Morningstar rating, and style is one of the nine Morningstar style ratings classifying a fund, which is a combination of size (small, mid, large) and investment style (value, blend, growth). The control variables are captured in the vector $X_{p,t}$ with γ being the corresponding coefficient vector. The control variables are lagged flows for month t – 19, net expense ratio, a dummy for no-load funds, five-year return volatility, log of fund size and log of fund age. Standard errors are also double clustered by fund and month to account for heteroskedasticity within these clusters.

The coefficients b_1, b_2, b_3, b_4 and b_5 are the parameters of interest and measure how flows respond to different return components. Rational investors assessing the performance of a fund manager would account for the known factors and only reward managers than can deliver alpha. In this case, the coefficients we would observe are $b_1 > 0$ and $b_2 = b_3 = b_4 = b_5 = 0$, as investors do respond to a funds four-factor alpha but do not respond to returns attributable to common factors. Less sophisticated investors might not observe that a funds return is partly attributable to exposure to common factors, in which case we would observe $b_i > 0$, $i=1,5$. It is important to note that the analysis uses fund flows, which by its nature is the aggregate flows of all investors of a fund. The coefficient estimates should thus be viewed as the flow response of the average investor on a particular factor-related return component.

Descriptive statistics

Table 1 provides descriptive statistics of the sample used in the main flow-return component regression. The main sample consists of 3,621 unique actively managed European mutual funds spanning from September 2001 to April 2021, resulting in 207,253 fund-month observations. Table 1, Panel A shows that the average fund has negative monthly flows of -0.10% during the sample period, which is similarly observed in the US data (Barber et al., 2016). The standard deviation of 9.32% indicates that the variation in fund flows is substantial. The average monthly fund return is positive at 0.72%, but an interquartile range of 6.06% and standard deviation of 5.56% show that there is considerable variability in returns across time and funds. The average fund has a size of 545 million USD, while the median fund is about one-third the size at 187 million USD. The average fund age is 16.8 years, while the interquartile range is only about 3.9 years. The average fund age is quite large, which is influenced by the data requirement of five-years of return data, tilting the sample towards mature funds. The average annual net expense ratio is 1.58% and about half of the fund-month observations (53%) do not have a front- or back-end load. The average 1- and 5-year volatility at 4.99% and 5.42% are very comparable to the standard deviation of fund returns across fund-month observations, which is to be expected.

Table 1

Descriptive statistics

	#Obs	Mean	SD	25th perc	50th perc	75th perc
Panel A: Fund characteristics						
Percentage fund flows	207,253	-0.10%	9.32%	-1.65%	-0.35%	0.76%
Fund return (monthly)	207,253	0.71%	5.56%	-2.19%	0.95%	3.86%
Fund size (mil USD)	207,253	545.530	1145.122	69.396	187.306	511.552
Fund age (months)	207,253	201.680	59.842	170.000	192.000	217.000
Expense ratio	207,253	1.584%	0.610%	1.240%	1.580%	1.890%
No load dummy	207,253	0.533	0.499	0.000	1.000	1.000
Volatility 1 year	207,253	4.994%	2.200%	3.394%	4.501%	6.135%
Volatility 5 year	207,253	5.416%	1.756%	4.057%	5.040%	6.446%
Star rating	207,253	3.291	1.007	3.000	3.000	4.000
Panel B: Fund alpha and factor exposures						
Alpha	239,537	-0.119%	0.421%	-0.364%	-0.139%	0.096%
Beta	239,537	1.057	0.189	0.953	1.059	1.171
Size coefficient	239,537	0.098	0.382	-0.164	0.041	0.308
Value coefficient	239,537	-0.077	0.367	-0.259	-0.067	0.123
Momentum coefficient	239,537	-0.004	0.189	-0.096	-0.001	0.088
Adj. R2	239,537	0.761	0.170	0.685	0.808	0.886
Panel C: Exponentially weighted return components						
ALPHA	207,253	-0.151%	1.071%	-0.659%	-0.159%	0.327%
MKTRET	207,253	0.721%	1.691%	0.235%	0.884%	1.616%
SIZRET	207,253	0.001%	0.216%	-0.065%	0.001%	0.074%
VALRET	207,253	0.007%	0.391%	-0.079%	0.010%	0.118%
MOMRET	207,253	-0.022%	0.287%	-0.073%	-0.004%	0.056%

Panel D: Mean descriptive statistics across Morningstar ratings

	1 Star	2 Star	3 Star	4 Star	5 Star
Percentage fund flows (avg)	-0.70%	-0.71%	-0.46%	0.16%	1.39%
Fund return (avg; monthly)	0.24%	0.45%	0.67%	0.83%	1.05%
Fund size (avg; mil USD)	211.783	324.579	423.426	664.725	1049.191
Fund age (avg;months)	193.716	202.055	206.425	199.928	193.393
Expense ratio (avg)	1.922%	1.778%	1.608%	1.469%	1.415%
Volatility 5 year (avg)	6.192%	5.702%	5.456%	5.264%	5.014%
Alpha (avg)	-0.439%	-0.267%	-0.146%	-0.035%	0.114%
Beta (avg)	1.105	1.080	1.061	1.047	1.022
Size coefficient (avg)	0.206	0.132	0.083	0.075	0.090
Value coefficient (avg)	-0.054	-0.053	-0.067	-0.098	-0.121
Momentum coefficient (avg)	-0.063	-0.021	-0.003	0.007	0.022
Fund-month observations	7,511	36,621	75,815	62,729	24,577

Table 1 (continued)

Descriptive statistics

Panel E: correlation table between fund return components

	ALPHA	MKTRET	SIZRET	VALRET	MOMRET
ALPHA	1				
MKTRET	-0.0343	1			
SIZRET	-0.0919	0.0870	1		
VALRET	-0.1850	-0.0348	0.0181	1	
MOMRET	-0.1718	-0.1332	-0.0578	-0.0838	1

Panel A presents descriptive statistics of the sample used in the flow-return component regression. These statistics are fund-month observations starting in September 2001 and ending in April 2021. Flows are estimated as the percentage growth of a funds TNA from month $t-1$ to t, adjusted for a fund returns in month t. No load dummy is true if none of the share classes of a fund has a front- or back-end load. One year volatility is measured as the standard deviation of a funds return over the prior months $t-1$ to $t-12$ and five-year volatility is measured over the prior months $t-1$ to $t-60$. Fund flows are trimmed below -90% and above 1,000%. Net expense ratio and 1- and 5-year volatility are winsorized at the 1^{st} and 99^{th} percentile.
Panel B presents the coefficients of the Carhart (1997) four-factor rolling regression across fund-month observations for the period starting in March 2000 and ending in April 2021.
Panel C presents descriptive statistics of exponentially weighted fund return components. For each fund-month observation, the return components are estimated by first multiplying a funds factor loading, estimated in Equation (2), times the factor returns. Then, each month t return component is estimated as the exponentially weighted average of the return component over the prior 18 months (see Section 'Decay function and model specification' for more details).
Panel D Presents average means for each star rating of variables and coefficients of the four-factor rolling regression, based on fund-month observations. These variables and coefficients are the same as defined in Panel A and panel B.
Panel E presents the correlation between exponentially weighted fund return components.

Table 1, Panel B presents coefficients of the rolling window four-factor regression specified in Equation (2). This includes the 18 months prior to the main sample, explaining the greater number of fund-month observations. Consistent with the vast documentation in the mutual fund literature (Carhart, 1997; Jensen, 1968; Malkiel, 1995), the average (European) mutual fund has a negative monthly alpha of -0.12% (-1.42% per year). The average market beta is 1.06, suggesting that the average fund follows the market portfolio closely. The average coefficients of size, value and momentum are considerably smaller at 0.10, -0.08 and -0.00, respectively. This suggest that the average fund has little exposure to these factors. However, there exist a substantial level of variation in factor loadings between funds, as the standard deviation for the size, value and momentum are much greater than the mean at 0.38, 0.37 and 0.19, respectively. In contrast, the standard deviation on the market beta is relatively modest at 0.19.

Descriptive statistics of factor-related return components, the independent variables in the main regression, are presented in Table 1, Panel C. The average exponentially weighted alpha return component is -0.15%, in line with the on average negative alpha from the four-factor regression. The market return component is 0.72% on average and thus is a large part of the average fund returns. The other factor return components are all considerably smaller on average. More interesting is once more the large variability in return components. With standard deviation of 1.07%, 1.69%, 0.22%, 0.39% and 0.29% for the alpha-, market-, value-, size- and momentum return components, respectively, being considerably larger than the mean.

As this research aims to discover whether the easy-to-access and -process signal of Morningstar ratings plays a role in European mutual fund investors investment decision, it is interesting to highlight the average mean descriptive statistics for each Morningstar rating. Table 1, Panel D has an intriguing story, showing that the average 1-star fund is quite different from the average 5-star fund. To start, the average 5-star funds earn 0.81% more returns per month (10.16% per year) than the average 1-star fund. The average 1-star fund also has 1.18% more volatility than the average 5-star fund, two-thirds of the standard deviation of the entire sample. These statistics are to be expected, as the Morningstar rating system rewards returns and penalizes high volatility by construction. More interesting to this research are the average flows to differently rated funds. The 5-star rated funds receives a substantial 2.09% more flows per month than 1-star funds on average. Furthermore, only 4- and 5-star rated funds receive positive flows. The magnitude of the difference in percentage and dollar flows across star ratings are visually represented in Appendix B. While these results show that ratings are associated with flows, they do not suggest that ratings drive these flows, since the ratings are highly correlated with sophisticated performance measures. As shown in Table 1, Panel D, only the 5-star rated funds have on average a positive alpha in the four-factor rolling window regression at 0.11%, above the 75th percentile of the entire sample. The results section will go into more detail in uncovering whether Morningstar ratings are the actual source of flows. Remarkably, the average 5-star fund is large, roughly five times the size of the average 1-star fund and double that of the average fund in the sample. This adds evidence to the mixed consensus in the literature on the diseconomies of scale in the mutual fund industry, consistent with some of the more recent researches pointing towards there being no diseconomies of scale (Phillips et al., 2018). Moreover, 5-star funds pay 0.51% less fees annually than 1-star funds, consistent with the negative relation between costs and performance, found by for instance Gil-Bazo and Ruiz-Verdú (2009) and Carhart (1997). The combination of larger funds performing better and charging smaller fees than smaller funds is also consistent with the findings of Elton et al. (2012). Finally, the factor loadings do not differ remarkably much between star ratings, although higher rated funds have moderately less exposure to the market-, size- and value coefficients and a slightly greater exposure to the momentum factor.

Before moving on to the results, it is relevant to establish that the independent variables in the main regression are not highly correlated amongst each other, which otherwise may impact the quality of the interpretation of the independent variables. The correlations between the independent variables of the main regression are presented in Table 1, Panel E. In general, the correlations between the factor-related return components are low, with the greatest being -0.19.

4. Results

Top-ranked performance

To set the stage, I investigate the flows received by top- and bottom ranked funds based on different performance measures. In each month, the data is sorted on a performance measure and then the number of 5-star funds in each month are used to classify top- and bottom ranked funds. For example, when a month has 100 5-star funds, the 100 funds with highest and lowest four-factor alpha are classified as top- and bottom ranked four-factor funds for that month, respectively[7]. This is ranking is done for the Morningstar ratings, market-adjusted returns, CAPM alpha, three-factor alpha and four-factor alpha performance measures. Of the observations in the sample, 11.9% are rated as 5-star funds, meaning that top- and bottom ranked funds of each measure both have 11.9% of the total sample observations. With these classifications, the average fraction of positive flows, the average percentage flows and the average dollar flows are computed to top- and bottom ranked funds, displayed in Table 2.

Table 2

Flows to top- and bottom ranked funds

	Positive flows (%)			Avg fund flow (%)			Avg fund flow ($ mil)		
	Top	**Bottom**	**Diff**	**Top**	**Bottom**	**Diff**	**Top**	**Bottom**	**Diff**
Morningstar	56.56	29.35	27.21	1.39	-0.80	2.19	3.01	-2.13	5.14
Market adjusted	44.72	34.37	10.35	0.63	-0.82	1.46	0.08	-2.75	2.84
CAPM	44.48	33.93	10.55	0.65	-0.83	1.48	-0.10	-2.72	2.62
Three-factor	43.48	34.80	8.69	0.55	-0.72	1.27	-0.29	-2.48	2.19
Four-factor	43.19	35.10	8.08	0.53	-0.71	1.24	-0.44	-2.45	2.01

This table presents fund-month averages of top- and bottom ranked funds. For each month, the number of 5-star funds (Nt) are counted. Then, the data is sorted based on a performance measure and the number of 5-star funds (Nt) are used to classify the top- and bottom ranked funds each month. The performance measures are Morningstar ratings, market adjusted return, CAPM alpha, three-factor alpha and four-factor alpha. Note that since the ranking occurs monthly, ranking on market adjust return results in the same ranking as ranking on unadjusted- or risk-free adjusted returns. Positive flows (%) is the fraction of positive flows going to top- and bottom ranked funds. Avg. fund flow (%) is the average percentage net flow going to top- and bottom ranked funds, and Avg. fund flow ($ mil) is the average flow in millions of USD going to top- and bottom ranked funds. The top- and bottom ranked funds both represent 11.9% of the entire sample.

[7] This is done to maintain comparability of dollar flows between top- and bottom ranked funds. The analysis is also performed with the number of 1-star funds in each month to classify top- and bottom ranked funds, and results remain qualitatively similar. The Table 2 results also remain qualitatively similar when the number of 5-star and 1-star funds is used to classify top- and bottom ranked, respectively.

Of the top Morningstar rated funds, 56.6% receive positive flows, while only 29.3% of the bottom ranked funds receive positive flows. These are both the highest and lowest fractions of positive flows for the top- and bottom ranked funds of any performance measure, respectively. The gap between the two is also the largest by a considerable margin at 27.2%, whereas the next largest gap is 10.6% for the CAPM alpha measure. A similar story holds for the average percentage- and dollar flows, where the top Morningstar rated funds receive significantly more flows than the other performance measures and the gap between top- and bottom Morningstar ranked funds is between 1.5 and 2.6 times the size of the gap of all the other performance measures. Relative to the Morningstar rating, all the other performance measures have results which are comparable to each other. Notably, the top ranked CAPM-, three-factor- and four-factor alpha funds receive a negative dollar flow on average, while the top ranked market-adjusted return funds receive only about 3% of the flows the top ranked Morningstar rating funds receive.

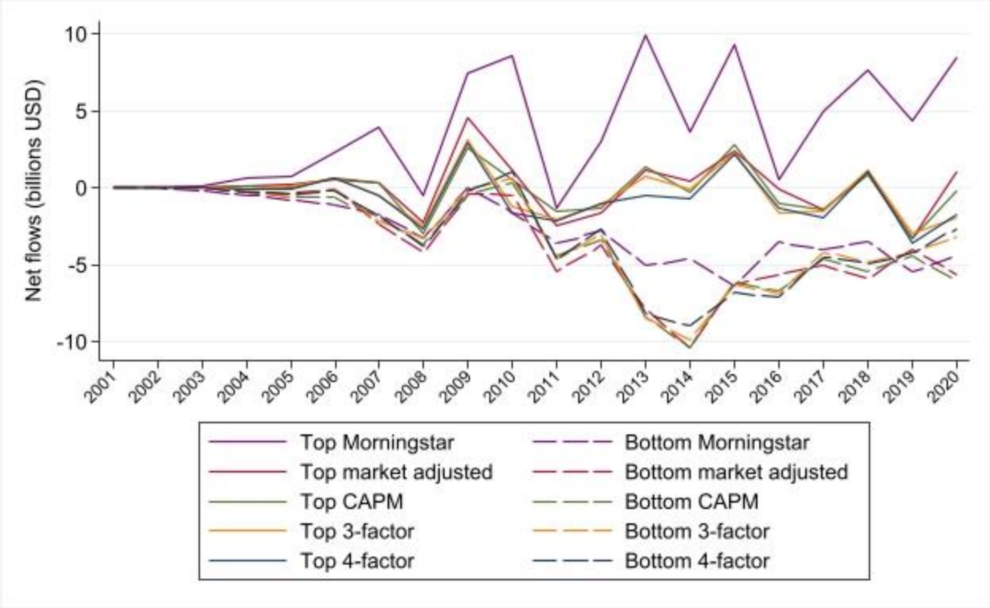
From this, it appears that Morningstar ratings outperform any other measure in predicting flows. To elaborate on the magnitude of this outperformance, Figure 2 presents the net flows to top- and bottom ranked funds of different performance measures. Figure 2, Panel A shows the annual net flows to the top- and bottom ranked funds of different performance measure. What is clear is that the top Morningstar ranked funds receive an economically and statistically significant amount more flows annually relative to top ranked funds of any other performance measures; About 3.6 billion USD more on average than the next best performance measure, which is market-adjusted returns. Figure 2, Panel B amplifies these results even further by plotting the accumulated flows to the top- and bottom ranked funds of each performance measure over the sample period. Top Morningstar rated funds clearly outclasses top-ranked fund based any other performance measure in accumulating flows. Strikingly, the market-adjusted return performance measure is the only performance measure where the top-ranked funds are able to accumulate positive flows over the sample period. When performance measures rank the worst funds, the difference in flows predicted by the performance measures are less pronounced, although it is worth pointing out that the bottom ranked CAPM alpha and market-adjusted return funds get punished the most.

What is apparent from these results is that 5-star rated funds receive substantially more capital from investors than the best performing funds ranked according to several different asset pricing model. The results thus indicate that Morningstar ratings are a strong predictor of mutual fund flows. Additionally, relative to Morningstar rating, the difference in the ability to predict flows by all the other performance measures is marginal. These results are remarkable, as performance measures such as the four-factor alpha, which is backed by financial theory, should do a much better job at explain the skill of a fund manager. This suggests that the average European mutual fund investor either naively follows Morningstar rating or is boundedly rational, outsourcing the complicated risk-adjustment performance

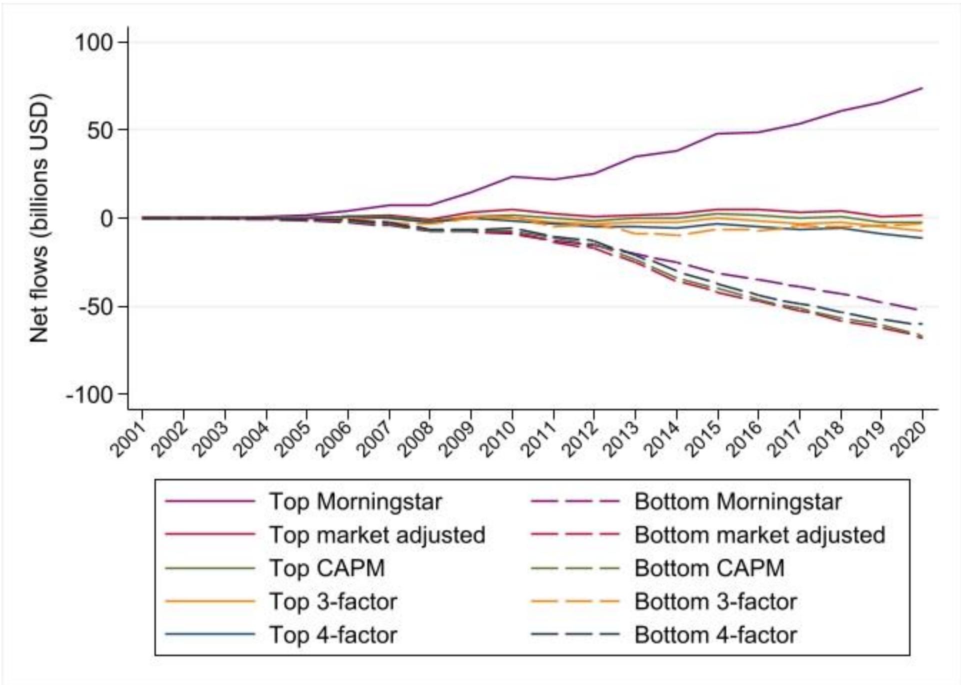
Figure 2

Flows to top- and bottom ranked funds

Panel A: Annual flows to top- and bottom ranked funds



Panel B: accumulated flows to top- and bottom ranked funds



These figures present the aggregate flows to top- and bottom ranked funds of various performance measures. For each month, the data is sorted on a performance measure and then the number of 5-star funds in each month are used to classify top- and bottom ranked funds. The performance measures include Morningstar ratings, market adjusted return, CAPM alpha, three-factor alpha and four-factor alpha. Panel A presents the annual aggregate dollar flows to top-and bottom ranked funds across the sample period. Panel B presents the dollar flows accumulated by top- and bottom ranked funds.

measurement to a third-party. Finally, it is interesting to point out the top ranked market-adjusted return funds receive significantly more flows than the top ranked three- and four-factor funds at conventional levels, as well as receiving more flows than the top ranked CAPM funds at the 10% level. This analysis thus gives the impression that European mutual fund investors on average care most about Morningstar ratings, and then market-adjusted returns. This is somewhat in contrast with the main panel regression analysis, which shows that investors do account for the market-related return component of a fund. A possible explanation for this may be that investors implicitly account for market risks when they are chasing market-adjusted returns.

Main analysis: Effect of factor-related return on fund flows

In the main analysis I recreate the panel regression of fund flows on alphas and factor-related return components, weighted over the prior 18 months, to infer what factors matter to European mutual fund investors when considering the performance of a mutual fund manager. Table 3 shows the results, where Column (1) is the main model with standard controls and month fixed effects. Column (2) is the second model which adds Morningstar style categories to create a two-way month-style fixed effect. Column (3) is the third model, which further adds Morningstar ratings to create a three-way month-style-rating fixed effect. Since the second and third model add multiple levels of fixed effects, it creates singleton groups, which have been removed from the regression. This causes slightly fewer observations in the second and third model, relative to the first model.

The first model shows that mutual fund flows respond positively to the four-factor alpha, with an estimated sensitivity of 0.78, significant at the 1% level. In terms of magnitude, a one standard deviation increase in alpha is associated with an 0.84% increase in monthly fund flows[8]. Contrasting this, the market-related return component has an insignificant effect on mutual fund flows. This suggest that *any* performance achieved by European mutual fund manager attributable to the market factor, is not rewarded by flows. European mutual fund investors thus appear to tend more to the market risk factor than U.S. investors do, as the coefficient found by Barber et al. (2016) is 0.25 and highly significant. Column (1) also shows that the other factor-related return components are all highly significant and positive. Mutual fund flows are even more sensitive to the value factor return component, and to a lesser degree the size-factor return component, than they are to the four-factor alpha, at 1.00 and 0.83, respectively. This does not hold, however, for the momentum factor with a coefficient of 0.64. The stronger and weaker sensitivity of the value- and momentum return components relative to the four-factor alpha may be consistent with Choi and Roberston (2020). They find in a survey of U.S. households that individuals generally hold the believe that the value factor is safer and does not earn higher expected returns, and that the momentum factor is riskier and does not earn higher expected returns. Funds that have a substantial value- or momentum related return component may

[8] 1.071%*0.7803 = 0.836%

Table 3

Results from panel regression of fund flows on alpha and factor-related return components

	(1)	(2)	(3)
ALPHA	0.780***	0.781***	0.741***
	(0.047)	(0.048)	(0.048)
MKTRET	-0.028	-0.030	-0.021
	(0.050)	(0.050)	(0.052)
SIZRET	0.830***	0.766***	0.720***
	(0.135)	(0.144)	(0.146)
VALRET	0.998***	0.909***	0.868***
	(0.106)	(0.112)	(0.113)
MOMRET	0.636***	0.561***	0.534***
	(0.122)	(0.130)	(0.135)
Month fixed effects	Yes	No	No
Month-style fixed effects	No	Yes	No
Month-style-rating FE	No	No	Yes
Controls	Yes	Yes	Yes
Observations	207,253	207,189	206,197
R2	0.014	0.021	0.047
Adj. R2	0.013	0.012	0.012

This table presents panel regression results of fund flows on alpha and factor-related return components of the four-factor model. Mutual fund flows (independent variable) are estimated as the percentage growth of a funds TNA from month t – 1 to t, adjusted for a fund returns in month t. The Carhart (1997) four-factor model includes the market-, size-, value- and momentum factor. For each fund-month observation, the return components are estimated by first multiplying a funds factor loading, estimated in Equation (2), times the factor returns. Then, each month t return component is estimated as the exponentially weighted average of the return component over the prior 18 months (see Section 2.3 for more details). The control variables are lagged flows for month t – 19, net expense ratio, a dummy for no-load funds, five-year return volatility, log of fund size and log of fund age. Factor return data is selected for developed market and retrieved from Kenneth French's Data Library, all other variables are retrieved from Morningstar Direct. Standard errors are double clustered by fund and month and are in parenthesis. *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.

advertise this, and retail investors may reward those funds according to the perceived level of risk and return. A different explanation for this effect comes from later robustness tests, which suggest that the value- and momentum coefficients are slightly upward and downward biased, respectively. Taken together, it seems that European mutual fund investors fully account for market risk of a fund when evaluating fund performance. However, investors do not appear to tend to the value-, size- and, to a lesser extent, the momentum exposure of a fund. At the same time, investors appear to react strongly to a fund's alpha.

The second model has month-style fixed effects in place of month fixed effects. The second model thus also accounts for unobserved heterogeneity across the nine Morningstar style categories. In general, the sensitivity of fund flows to each return component decreases slightly, while significance of the coefficients remains the same to the previous model specification. For instance, the sensitivity of fund flows to the value-return component decreases by 8.9% to 0.91. This is consistent with Barber et al. (2016), which show that as style categories explain some of the variation in value- and size tilts of funds it can slightly dampen the sensitivity, but at the same time investors generally treat any returns attributable to style categories as alpha. One way to interpret this is that the potential style chasing behaviour of investors does not explain these results. Another way of interpreting these results, is that within each Morningstar style category, the results and the inference of the first model still holds.

In the third model, the month-style fixed effects are replaced by month-style-rating fixed effects. Remember that Morningstar ratings by construction reward high return and low volatility funds relative to other funds within their peer group. Hence, by following Morningstar ratings, it may be possible that investors account for market risk and potentially even other factors when evaluating fund managers performance. The results suggest that this is not the case. Column (3) present the results. The coefficients for each return component decreased between 5% and 16% relative to coefficients of the first model. While the high correlation between Morningstar ratings and performance decreases the sensitivity of flows to return components moderately, the results remain qualitatively similar. Morningstar ratings thus appear not able to explain the results. In other words, within each Morningstar rating, European mutual fund investors appear to fully account for market risk when assessing fund performance, but do not consider the value-, size- and, to a lesser extent, the momentum exposure of a fund.

Time varying FPS and robustness of results

Concerns have been raised in the recent literature about the evidence found in favour of the CAPM model in the panel regression with time-fixed effects. Specifically, Ben-David et al. (2019) argues that the time-varying flow-to performance sensitivity (FPS), first documented by Franzoni and Schmalz (2017), leads to spurious evidence which would give the impression that investors account for market risk by using the CAPM model to assess mutual fund manager's performance. The following sections will address this concern as follows. First, the concern raised by Ben-David et al. (2019) will be described in detail. Second, it will be investigated whether the European mutual market exhibits the time-varying FPS, documented by Franzoni and Schmalz (2017) for the U.S. mutual fund market. Third, it will be tested if the dispersion in the market- and other factor-related return components is large in periods when the FPS is weak, causing downward biased coefficients. Fourth, the main panel regression will be performed again using simulated flows, which by construction do not differentiate between return components. Performing the panel regression using simulated- and actual flows leads to nearly identical results in the U.S. mutual fund sample used by Barber et al. (2016), suggesting that the main results are found mechanically. To preview the results, this is not the case in the European mutual fund data, which suggest that the results found in the panel regression of this research are likely not spurious.

Why the results might be spurious

Franzoni and Schmalz (2017) document that the cross-sectional FPS is a hump-shaped function of aggregate market risk. In times of moderate market returns the FPS is twice or even thrice as large as in extreme market return. Put differently, during times in the market with either very high or low returns, fund flows respond less to past performance. At those same times of extreme market return, the cross-sectional dispersion in market-related fund returns, MKTRET, is particularly large. This is by construction, as MKTRET is computed as a fund's beta (which does not vary greatly over time) times the markets realized return. Combined, in periods where the FPS is weak the market-related return is more dispersed across funds. This observation is key when combined with the following econometric fact. In a panel regression with time-fixed effects, the coefficients are largely influenced by those observations in volatile periods. Pastor et al. (2017) show that in a panel regression with time-fixed effects the coefficient estimates are the weighted average across periods of period-by-period cross-sectional regressions coefficients. Specifically, coefficient estimate $\hat{b}_x = \sum_{t=1}^{T} w_t \hat{b}_{x,t}$ where each periods cross-sectional coefficient estimate $\hat{b}_{x,t}$ is weighted by w_t. The weight w_t is proportional to each periods cross-sectional variance of the independent variable times the number of observations in each period[9]. Put simply, the periods in which flows are the least sensitive to performance determine the coefficient estimates for a large part. Even if investors did not distinguish between market-related returns and the alpha of a fund, the market-related return coefficient is likely downward biased relative to the alpha coefficient estimate, since the market-related return component is likely more dispersed than the alpha return component, by its construction.

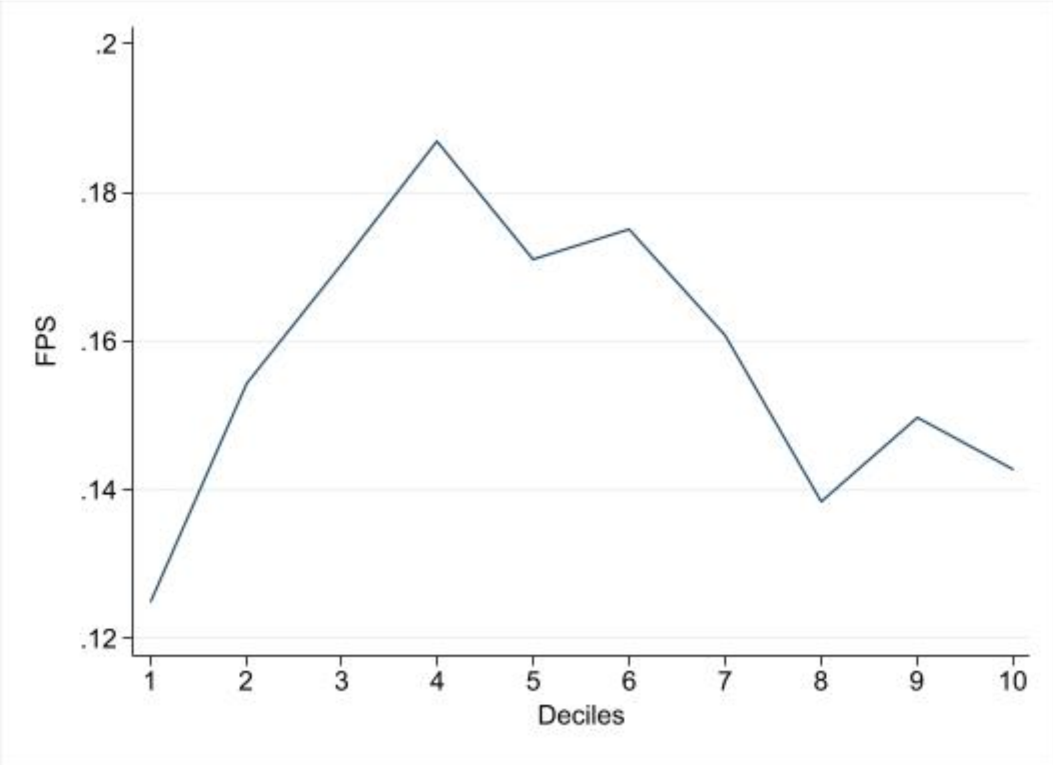
The time-varying flow-performance sensitivity of the European fund market

To test whether this concern applies to the European mutual fund market, it must first be established that this market exhibits a time-varying FPS. Following Franzoni and Schmalz (2017), the FPS is estimated as the slope in a monthly cross-sectional simple regression of flows on the past quarter of total fund returns. Then, the 232 months in the sample are sorted into to ten buckets based on past quarter total excess market returns. Figure 3 presents the average FPS in each bucket. It shows that the European mutual fund sample exhibits a hump-shaped relation between FPS and market

[9] Formally, $w_t = \left(N_t \hat{\sigma}_{x,t}^2\right) / \left(\sum_{t=1}^{T} N_t \hat{\sigma}_{x,t}^2\right)$ where N_t denotes the total observations in each period *t* and $\hat{\sigma}_{x,t}^2$ denotes the cross-sectional variance of the independent variable *x* in each period *t*.

Figure 3

Flow-performance sensitivity on deciles of market realizations



This figure presents the average flow-performance sensitivity (FPS) on deciles of past quarter realized excess market return, in the same periods when the fund performance is measured. The FPS is estimated as the slope in a monthly cross-sectional regression of flows on past quarter total fund returns. The deciles sort the 232 months in the sample based on past quarter total excess market returns. The lines represent the average FPS in that decile.

realizations, although it is much less pronounced than is documented by Franzoni and Schmalz (2017). In moderate periods of market returns, the FPS is at most 1.5 times the size of the lowest decile of extremely low market realizations, and 1.3 times the size of the highest deciles of extremely high market realizations. Since the difference in FPS between market states is less severe than that what is found in the U.S. fund market, the FPS is not as weak in those periods with high market return dispersion. In other words, the concern that those periods with weaker flow-performance sensitivity bias the coefficients downward should be less of an issue in this European mutual fund sample.

Return component dispersion during weak flow-performance sensitivity periods

In this next step, it will be investigated whether the market-related return component is indeed more dispersed in periods with weak FPS. Following Ben-David et al. (2019), for each month, a simple cross-sectional regression of fund flows on the prior 18-month weighted total fund returns is performed, where the slope is defined as the FPS. Then, the 232 months are sorted into deciles based on the FPS. In Figure 3, each decile presents the average FPS across that decile in bars. For each decile, the average cross-sectional variance is computed for the total return over the prior 18 months and the fund return components (independent variables of interest). For the sake of discussion, the lines in Figure 3 present the average cross-sectional variance in each decile relative to the total average variance across all deciles. There a total of five plots, in each the four-factor alpha dispersion is displayed for comparability.

Figure 4, Panel A shows the average total return dispersion and the average alpha dispersion across FPS deciles. The total return dispersion is indeed greater in the first few weak FPS deciles than in the strong FPS deciles. From the first through the fourth FPS decile, the total return dispersion is greater than the four-factor alpha dispersion. However, the difference in total return dispersion between the bottom- and top FPS decile is small relative to what was found by Ben-David et al. (2019), They find that the cross-sectional total return dispersion in the highest FPS decile is about four times as large as in the lowest FPS decile, whereas in the European mutual fund sample it is only about 1.4 times as large. These results are consistent with the previous exercise, which also showed that the market-state dependent flow-performance sensitivity phenomenon is weaker in the European mutual fund market than the U.S. market.

More important to this research is the dispersion in market-related return. Figure 4, Panel B shows the cross-sectional dispersion of each factor-related return per decile. In stark contrast to Ben-David et al. (2019), the dispersion in market-related returns does not appear to be particularly high when the FPS is weak. The market-related return dispersion is actually slightly greater in the top FPS decile than in the bottom FPS decile. This is remarkable, considering that Ben-David et al (2019) find that the

Figure 4

Dispersion in total return and return components on deciles of flow-performance sensitivity

Panel A: Cross-sectional dispersion of total fund returns and alphas

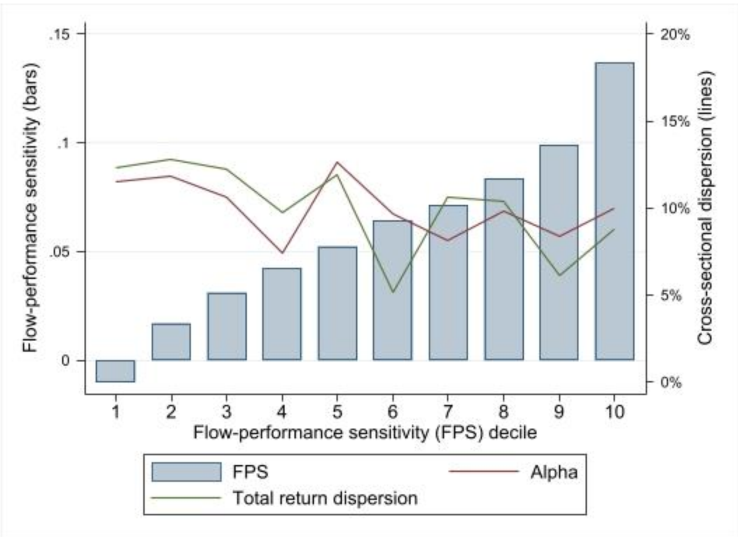
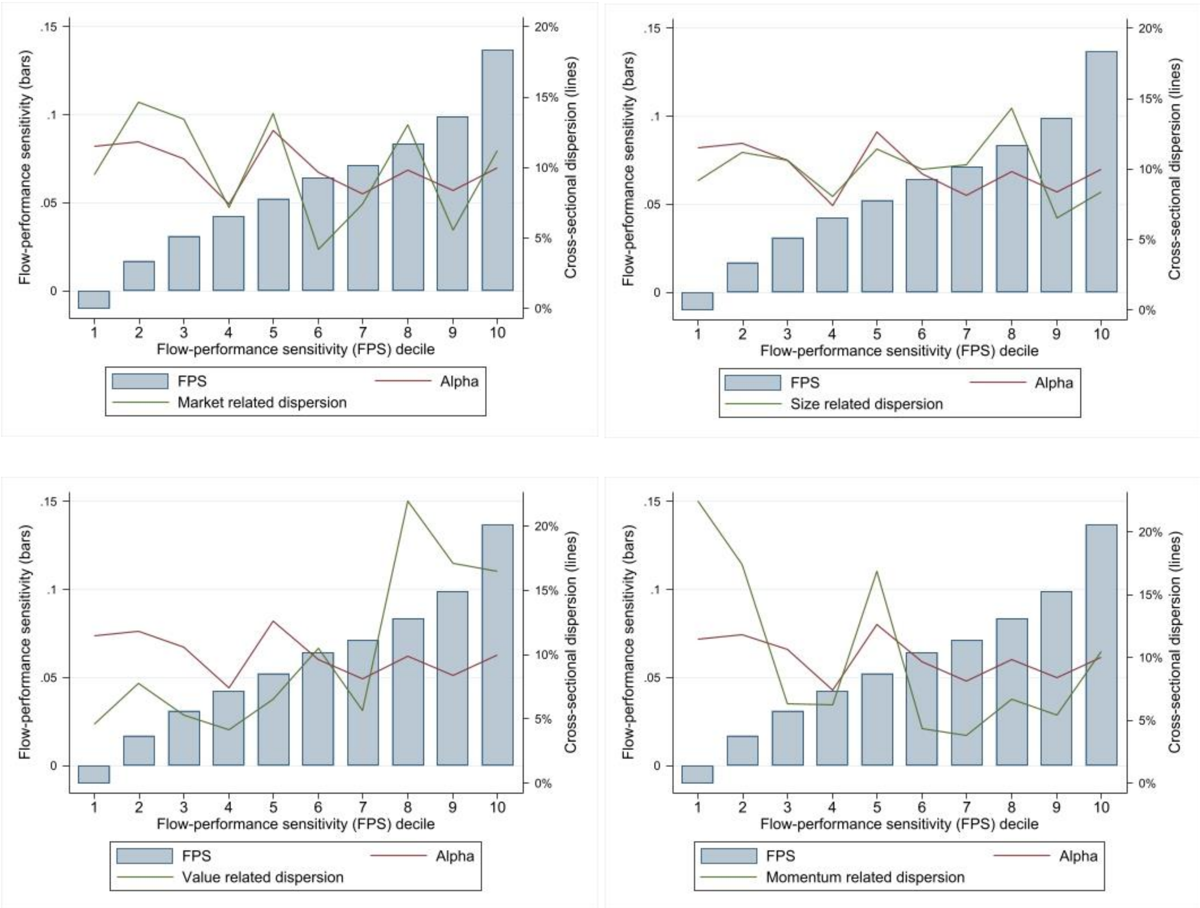


Figure 4 (continued)

Dispersion in total return and return components on deciles of flow-performance sensitivity

Panel B: Cross-sectional dispersion of factor-related return components



These figures show the cross-sectional dispersion in total fund past 18-month fund returns and return components across deciles of flow-performance sensitivity. The bars present the average FPS in that decile, where FPS is measured as the slope in a monthly cross-sectional regression of flows on past 18 month weighted total fund returns. The return components are the same as the independent variables used in the main panel regression, Equation (8), which are MKTRET, VALRET, SIZRET, MOMRET, ALPHA. The alpha return component is shown in each graph for comparability. The lines represent the average cross-sectional variance in each decile divided by the total average variance across all deciles.

market-related return dispersion is about eight times as large in the bottom FPS decile relative to the top decile. This indicates that there likely is not a substantial bias in the results found in this research, as the panel regression coefficients estimates are not largely influenced by those periods where the FPS is weak.

A similar story holds for the size-related return component, as in each FPS decile its cross-sectional dispersion is not excessively greater than the other FPS deciles. Remarkably, the dispersion in value-related returns shows the opposite pattern that was found by Ben-David et al. (2019), where the dispersion is the greatest in periods when the FPS is strongest. Since the dispersion in the strong FPS decile is about three to four times larger than the dispersion in the weak FPS decile, this may suggest that the size-related return coefficient estimate is biased upward. This could explain why flows appear to be more sensitive to the size-related return component than the alpha of a fund in the main panel regression. Only the dispersion in momentum-related returns appears to be somewhat consistent with Ben-David et al. (2019). The dispersion is greater when the FPS is weak and smaller when the FPS is strong, although it is again of a much smaller magnitude found by Ben-David et al. (2019)[10]. In general, it appears that the dispersion in total return and other factor related components is not severely larger in the periods of weak FPS than in those of strong FPS. This is especially the case for the main variable of interest, the market-related return component. Hence, the regression coefficient estimate of the market-related return component is likely not substantially biased downward. This analysis does, however, provide moderate evidence that the value- and momentum related return component coefficient estimates are biased upwards and downwards, respectively.

Sensitivity of simulated flows to fund return components

To formally test whether the coefficient estimates are biased, I follow Ben-David et al (2019) and simulate flows, which by construction do not differentiate between return components. These are subsequently used as the dependent variable in the main panel regression to discover whether the results were found mechanically. To simulate flows, for each month *t*, cross-sectional regressions are run of fund flows on unadjusted realized fund returns and controls. The flows that these models predict are the flows that would be observed under the null hypothesis that investors only care about unadjusted returns. Two models are used, the first is specified as follows

$$F_{p,t} = b_0 + \gamma X_{p,t} + b_1 R_{p,t} + e_t \tag{9}$$

Where $R_{p,t}$ is the past 18-month weighted total return of fund p and $X_{p,t}$ is the same vector of control variables used in the main panel regression, defined in Equation (8). The second model also adds Morningstar ratings and is specified as follows

$$F_{p,t} = \gamma X_{p,t} + \sum_{k=1}^{5} \gamma_{p,t}^{k} \, \mathrm{I}_{(star=k)} + b_1 R_{p,t} + e_t \tag{10}$$

[10] In this sample, the momentum-related return dispersion in the lowest FPS decile is about twice the size of that of the highest FPS decile, whereas in the sample of Ben-David et al. (2019) it appears to be more than twenty times as large.

Table 4

Response of real- and simulated flows on return components

	Observed flows	Observed flows under null hypothesis	
	(1)	(2)	(3)
ALPHA	0.780***	0.471***	0.450***
	0.047	0.020	0.021
MKTRET	-(0.028)	0.051**	0.050**
	0.050	0.025	0.025
SIZRET	0.830***	0.564***	0.546***
	(0.135)	(0.101)	(0.094)
VALRET	0.998***	0.708***	0.675***
	(0.106)	(0.052)	(0.054)
MOMRET	0.636***	0.437***	0.410***
	0.1215	0.08257	0.08315
Month fixed effects	Yes	Yes	Yes
Month-style fixed effects	No	No	No
Month-style-rating fixed effects	No	No	No
Controls	Yes	Yes	Yes
Observations	207,253	207,253	207,253
Adj. R2	0.013	0.010	0.010

This table presents panel regression results of real- and simulated fund flows on alpha and factor-related return components of the four-factor model. Flows are simulated under the null hypothesis that investors respond only to unadjusted total returns. Column (1) presents the models using actual flows. Column (2) presents simulated flows which only respond to cross-sectional variation in total fund returns. Column (3) presents results using simulated flows, which respond to Morningstar ratings in addition to total fund returns. Simulating flows involve predicting flows under the null hypothesis and bootstrapping residuals to get to simulated flows. The main panel regression is then re-estimated over 1,000 iterations using newly simulated flows each time. The control variables are the same as specified in Equation (8). Standard errors are double clustered by fund and month. Stander errors for Column (1) are panel regression standard errors, for Column (2) and Column (3) they are bootstrapped stander errors, all standard errors are presented in parenthesis. *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.

Where $\mathrm{I}_{(star=k)}$ is a set of dummy variable indicating whether a fund p has star rating k. This model is likely more realistic than Equation (9), as the analysis of flows to top- and bottom ranked funds shows that Morningstar ratings have a strong predictive ability for mutual fund flows. For each month t, the bootstrapped fund flow residuals are added to the predicted flows to obtain the simulated flows. The main panel regression is then performed again in 1,000 simulations, using differently simulated flows each iteration as the dependent variable instead. The results are presented in Table 4, where the coefficients and the empirical standard error reported are the average and the standard deviation of each coefficient across 1,000 simulations, respectively. Regressing on simulated flows causes the coefficients of the alpha-, size-, value- and momentum-related return component to decrease substantially. The market-related return component goes from having an insignificant effect on real flows, to a significant effect on simulated flows on conventional levels. However, the sensitivity is rather weaker relative to the sensitivity of other return components. These results are, in general, in contrast to the results found by Ben-David et al. (2019), as they find no statistical difference between the coefficient estimates.

Concluding this section, the results of these tests in general suggest that the coefficient estimates in the main panel regression are not found *mechanically*. However, as the sensitivity of the market-related return component is weak to simulated flows generated under the null hypothesis that investors only respond to unadjusted returns, flows may not be completely insensitive to market returns. Still, given that the dispersion in the market-related return component is not particularly large when the FPS is weak and that the coefficient estimate becomes significant at conventional levels using simulated flows, the bias is much less severe than proposed by Ben-David et al. (2019). A case can also be made that the value-related return component is upward biased, and that the momentum-related return component is downward biased, as the dispersion in those return components are large when the FPS is strong and weak, respectively. All things considered, the general inference of the main panel regression still holds; European mutual fund investors tend most to market risk when assessing fund performance, whilst generally considering returns attributable to common factors as a sign of outperformance.

Robustness of lambda

To ensure the results in the panel regression of fund flows on return components are not driven by the chosen decay rate parameter $\hat{\lambda}$, the panel regression is performed for a range of λ values. To refresh, the decay rate empirically found by Barber et al. (2016) is $\hat{\lambda}$ = 0.20551497, which is used to weight past mutual fund return components, such as specified in Equation (6). When λ = 0, each lagged return component is weighted equally. In this robustness test, the decay rate is changed by +/-50%, after which the panel regression specified in Equation (8) is performed again. This range is well within the range of reasonable decay rate values for this sample. For simplicity, only the first model is shown, which is the main model with month fixed effects. The results remain qualitatively similar across all model specification, as shown in Appendix C.

Shown in Table 5, the coefficient estimates increase as the decay rate $\hat{\lambda}$ decreases, and vice versa. Intuitively, as the lagged return components get weighted more equally across time, the sensitivity of the return components increases (and vice versa). In terms of magnitude, consider the coefficients estimate of the four-factor alpha return component; increasing (decreasing) $\hat{\lambda}$ by 50%

changes the coefficients estimate by 0.10 (-0.07). Remember that this range is likely excessively large compared to the true empirical λ. When increasing (decreasing) $\hat{\lambda}$ by a rather more reasonable 25%, the alpha coefficients change by 0.04 (-0.05). Taken together, the decay rate parameter λ influences the sensitivity of the return components on flows slightly, but the results remain qualitatively similar as the significance of coefficient estimates remains constant across λ paramaters. In conclusion, the specification of the λ parameter does not drive the results found in this research.

Table 5

Robustness test of the decay rate on the flow-return component regression

	Model 1				
	$\hat{\lambda}$	$\hat{\lambda}$ *50%	$\hat{\lambda}$ *75%	$\hat{\lambda}$ *125%	$\hat{\lambda}$ *150%
	(1)	(2)	(3)	(4)	(5)
ALPHA	0.780***	0.851***	0.825***	0.732***	0.685***
	(0.047)	(0.056)	(0.052)	(0.044)	(0.041)
MKTRET	-0.028	-0.015	-0.022	-0.032	-0.035
	(0.050)	(0.044)	(0.047)	(0.052)	(0.053)
SIZRET	0.830***	0.885***	0.865***	0.791***	0.752***
	(0.135)	(0.152)	(0.144)	(0.127)	(0.120)
VALRET	0.998***	1.084***	1.046***	0.947***	0.898***
	(0.106)	(0.122)	(0.113)	(0.101)	(0.096)
MOMRET	0.636***	0.649***	0.649***	0.617***	0.595***
	(0.122)	(0.142)	(0.132)	(0.112)	(0.103)
Month fixed effects	Yes	Yes	Yes	Yes	Yes
Month-style fixed effects	No	No	No	No	No
Month-style-rating FE	No	No	No	No	No
Controls	Yes	Yes	Yes	Yes	Yes
Observations	207,253	207,253	207,253	207,253	207,253
R2	0.014	0.013	0.014	0.014	0.014
Adj. R2	0.013	0.012	0.013	0.013	0.013

This table presents robustness test of the panel regression results of fund flows on alpha and factor-related return components from the four-factor model, using different decay rates to estimate the independent variables. The model presented here is the main model, which uses month-fixed effects. Column (1) presents the regression results when the decay rate $\hat{\lambda}$ = 0.20551497, the decay rate used in the main analysis. Columns (2), (3), (4) and (5) use $\hat{\lambda}$ times 50%, 75%, 125% and 150%, respectively. The control variables are the same as specified in Equation (8). Standard errors are double clustered by fund and month and are in parenthesis. *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.

5. Conclusion

This thesis set out to investigate what factors matters to European mutual fund investors when evaluating a fund managers performance. This adds to a recent body of literature, which provides mixed evidence for the U.S. market. Barber et al. (2016) find that investors tend most to the market risk factor, mostly treating other factor-related return components as alpha. Ben-David et al. (2019) re-examine their results and find that these results lead to spurious evidence in favour of investors using the CAPM. Mutual fund flows are instead most consistent with investors blindly following Morningstar ratings.

The empirical analysis of this thesis shows that European mutual fund investors fully account for the market risk factor when evaluating a fund's performance, caring even more about market risk than U.S. investors do. The results also show that European mutual fund investors do not attend to the value- and size related return components; flows are as sensitive to these return components as to a funds four-factor alpha. Investors appear to tend moderately to the momentum-related return component, which could be explained by household investors perception that momentum is risky and not rewarded with any returns (Choi & Robertson, 2020). Adding Morningstar ratings as a fixed effect does not explain these results, meaning that within each star rating these results still apply. However, in a test of flows to top- and bottom ranked funds, Morningstar ratings far outperform any other asset pricing model. Top Morningstar rated funds receive on average 3.6 billion USD more flows annually than the top ranked market-adjusted return funds. This surprisingly, is the next best performance measure in predicting flows, and those top ranked funds receive significantly more flows than the top ranked three- and four-factor alpha funds. A natural question that arises is why do investors appear to care about market risk, whilst top Morningstar rated funds garner the most flows? After all, if investors are capable of estimating a CAPM alpha, then there is not much preventing them from estimating a multi-factor alpha, other than the lack of knowledge about common factors explaining cross-sectional variation in returns. One reason might be that markets returns are a salient measure which is reported universally. Investors that pay attention to market returns when assessing a fund managers performance, may inadvertently appear to care about market risk. This conjecture would be consistent with top ranked market-adjusted return funds receiving comparable flows to top ranked CAPM alpha funds.

Concerns for the potential of spurious results are tested for. First, it is established that the European mutual fund market exhibits a time-varying flow-performance sensitivity, but this effect is much less pronounced than what is found in the U.S.. Part of the concern is that in a panel regression with time-fixed effect, the coefficients are largely influenced by those observations in volatile periods. In this sample, the dispersion in the market-related return component is not particularly great when the FPS is weak, meaning that the panel regression coefficient estimates are *not* largely influenced by those periods when the FPS is weak, which otherwise could downward bias the coefficient. This is formally tested for using simulated flows which would be observed under the null hypothesis that investors only care about unadjusted returns. The results are generally in contrast to what was found by Ben-David et al. (2019), suggesting that the coefficient estimates in the main panel regression are not found *mechanically* as simulated flows respond differently to return components than actual flows do. There is a case to be made that the market- and momentum related return components are somewhat downward bias, although this bias is far weaker than what is found for U.S. investors.

Taken together, it can be said, with a greater degree of certainty than can be said in the U.S. market, that European mutual fund investors account for market risk, whilst treating other factor-related return components as a sign of outperformance. While Morningstar ratings do not explain this effect, European mutual fund investors primarily let their investment decision be influenced by Morningstar ratings. Since the main variable of interest is aggregate fund flows, the results apply to the average investor. Considering this, it is probable that a large portion of investors are either boundedly rational or irrational and influenced by attention grabbing ratings, a smaller portion use a basic CAPM model to assess performance, and fewer still that use more sophisticated models. In aggregate, however, European mutual fund investors in their investment allocation decision use less sophisticated models and appear to be less than fully rational by using attention grabbing ratings.

Policy makers should consider making information about empirically backed performance measures more readily available to investors. Besides focusing on improving the financial literacy of households, regulators could instead focus on turning performance measures backed by sophisticated financial theory into easy-to-process and attention-grabbing metrics. This would be in the best interest of the general public, as fund managers in current state are incentivized to abuse this sub-optimal behavior of individual investors. It should be noted, however, that his study has a few potential limitations. The sample is inherently biased towards larger and mature funds, as five-years of prior returns are needed to estimate factor-loading. Additionally, due to the complexity and the limited time to investigate, this study did not parsimoniously model the decay rate. This issue does not affect the overall results qualitatively but does affect the sensitivity of the regression coefficients slightly. Motivated by the different results found in this study relative to the U.S. studies, future researchers might be interested to investigate what factors Asia-Pacific mutual fund investors care about.

Bibliography

Agarwal, V., Green, T. C., & Ren, H. (2018). Alpha or beta in the eye of the beholder: What drives hedge fund flows? *Journal of Financial Economics*, *127*(3), 417–434. https://doi.org/10.1016/j.jfineco.2018.01.006

Banegas, A., Gillen, B., Timmermann, A., & Wermers, R. (2013). The cross section of conditional mutual fund performance in European stock markets. *Journal of Financial Economics*, *108*(3), 699–726. https://doi.org/10.1016/j.jfineco.2013.01.008

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. https://doi.org/10.1093/rfs/hhw054

Barberis, N., & Thaler, R. (2005). Chapter 1. A Survey of Behavioral Finance. In *Advances in Behavioral Finance, Volume II* (pp. 1–76). Princeton University Press. https://www.degruyter.com/document/doi/10.1515/9781400829125-004/html

Ben-David, I., Li, J., Rossi, A., & Song, Y. (2019). *What Do Mutual Fund Investors Really Care About?* (SSRN Scholarly Paper ID 3292317). Social Science Research Network. https://papers.ssrn.com/abstract=3292317

Berk, J. B., & van Binsbergen, J. H. (2016). Assessing asset pricing models using revealed preference. *Journal of Financial Economics*, *119*(1), 1–23. https://doi.org/10.1016/j.jfineco.2015.08.010

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Chevalier, J., & Ellison, G. (1997). Risk Taking by Mutual Funds as a Response to Incentives. *Journal of Political Economy*, *105*(6), 1167–1200. https://doi.org/10.1086/516389

Choi, J. J., & Robertson, A. Z. (2020). What Matters to Individual Investors? Evidence from the Horse’s Mouth. *The Journal of Finance*, *75*(4), 1965–2020. https://doi.org/10.1111/jofi.12895

Clifford, C. P., Fulkerson, J. A., Jordan, B. D., & Waldman, S. (2013). *Risk and Fund Flows* (SSRN Scholarly Paper ID 1752362). Social Science Research Network. https://doi.org/10.2139/ssrn.1752362

Coval, J., & Stafford, E. (2007). Asset fire sales (and purchases) in equity markets. *Journal of Financial Economics*, *86*(2), 479–512. https://doi.org/10.1016/j.jfineco.2006.09.007

Del Guercio, D., & Tkac, P. A. (2008). *Star Power: The Effect of Morningstar Ratings on Mutual Fund Flow* (SSRN Scholarly Paper ID 286157). Social Science Research Network. https://doi.org/10.2139/ssrn.286157

Elton, E. J., Gruber, M. J., & Blake, C. R. (2012). Does Mutual Fund Size Matter? The Relationship Between Size and Performance. *The Review of Asset Pricing Studies*, *2*(1), 31–55. https://doi.org/10.1093/rapstu/ras001

Elton, E. J., Gruber, M. J., & Busse, J. A. (2002). *Are Investors Rational? Choices Among Index Funds* (SSRN Scholarly Paper ID 1300816). Social Science Research Network. https://papers.ssrn.com/abstract=1300816

Evans, R. B. (2010). Mutual Fund Incubation. *The Journal of Finance*, *65*(4), 1581–1611. https://doi.org/10.1111/j.1540-6261.2010.01579.x

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. https://doi.org/10.1111/j.1540-6261.2010.01598.x

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405X(93)90023-5

Franzoni, F., & Schmalz, M. C. (2017). Fund Flows and Market States. *The Review of Financial Studies*, *30*(8), 2621–2673. https://doi.org/10.1093/rfs/hhx015

Frazzini, A., & Lamont, O. A. (2008). Dumb money: Mutual fund flows and the cross-section of stock returns. *Journal of Financial Economics*, *88*(2), 299–322. https://doi.org/10.1016/j.jfineco.2007.07.001

Gil-Bazo, J., & Ruiz-Verdú, P. (2009). The Relation between Price and Performance in the Mutual Fund Industry. *The Journal of Finance*, *64*(5), 2153–2183. https://doi.org/10.1111/j.1540-6261.2009.01497.x

Hou, K., Xue, C., & Zhang, L. (2017). *A Comparison of New Factor Models* (SSRN Scholarly Paper ID 2520929). Social Science Research Network. https://doi.org/10.2139/ssrn.2520929

ICI. (2021). *2021 Investment Company Fact Book*. https://www.ici.org/system/files/2021-05/2021_factbook.pdf

Jegadeesh, N., & Mangipudi, C. S. (2021). What Do Fund Flows Reveal about Asset Pricing Models and Investor Sophistication? *The Review of Financial Studies*, *34*(1), 108–148. https://doi.org/10.1093/rfs/hhaa045

Jensen, M. C. (1968). The Performance of Mutual Funds in the Period 1945–1964. *The Journal of Finance*, *23*(2), 389–416. https://doi.org/10.1111/j.1540-6261.1968.tb00815.x

Malkiel, B. G. (1995). Returns from Investing in Equity Mutual Funds 1971 to 1991. *The Journal of Finance*, *50*(2), 549–572. https://doi.org/10.1111/j.1540-6261.1995.tb04795.x

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2017). Do Funds Make More When They Trade More? *The Journal of Finance*, *72*(4), 1483–1528. https://doi.org/10.1111/jofi.12509

Phillips, B., Pukthuanthong, K., & Rau, P. R. (2018). Size does not matter: Diseconomies of scale in the mutual fund industry revisited. *Journal of Banking & Finance*, *88*, 357–365. https://doi.org/10.1016/j.jbankfin.2017.12.005

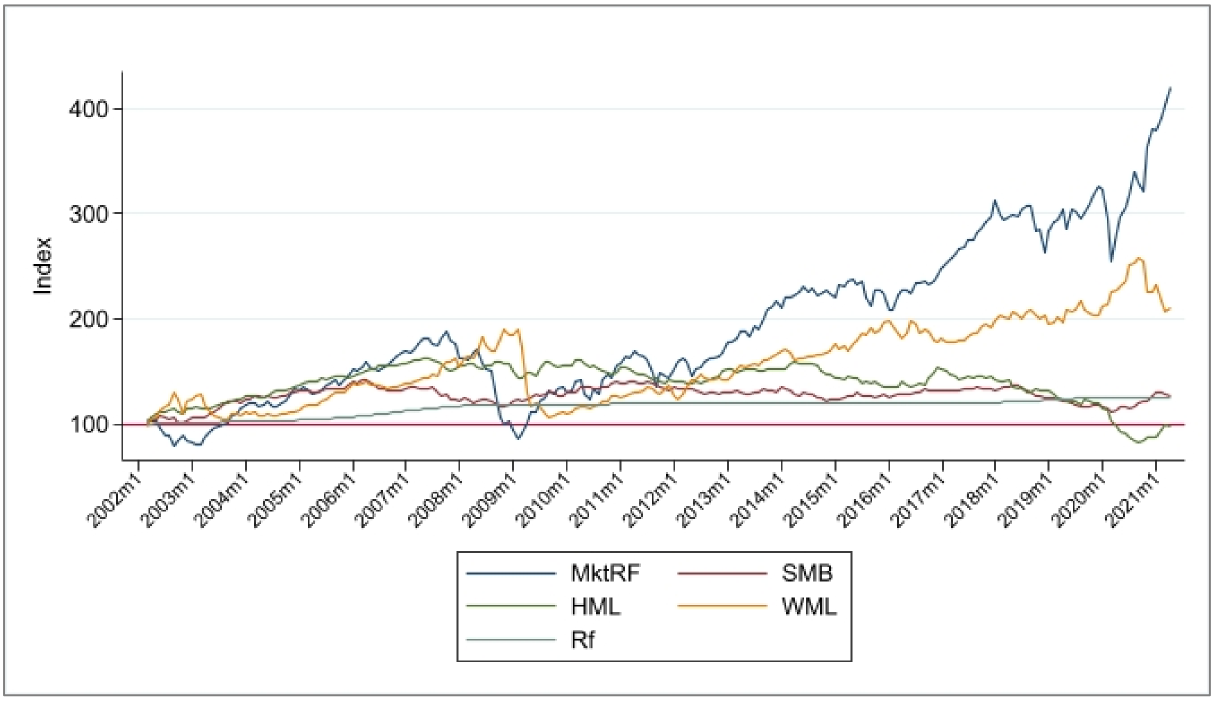
Sensoy, B. A. (2009). Performance evaluation and self-designated benchmark indexes in the mutual fund industry. *Journal of Financial Economics*, *92*(1), 25–39. https://doi.org/10.1016/j.jfineco.2008.02.011

Sirri, E. R., & Tufano, P. (1998). Costly Search and Mutual Fund Flows. *The Journal of Finance*, *53*(5), 1589–1622. https://doi.org/10.1111/0022-1082.00066

Solomon, D. H., Soltes, E., & Sosyura, D. (2014). Winners in the spotlight: Media coverage of fund holdings as a driver of flows. *Journal of Financial Economics*, *113*(1), 53–72. https://doi.org/10.1016/j.jfineco.2014.02.009

Appendix A

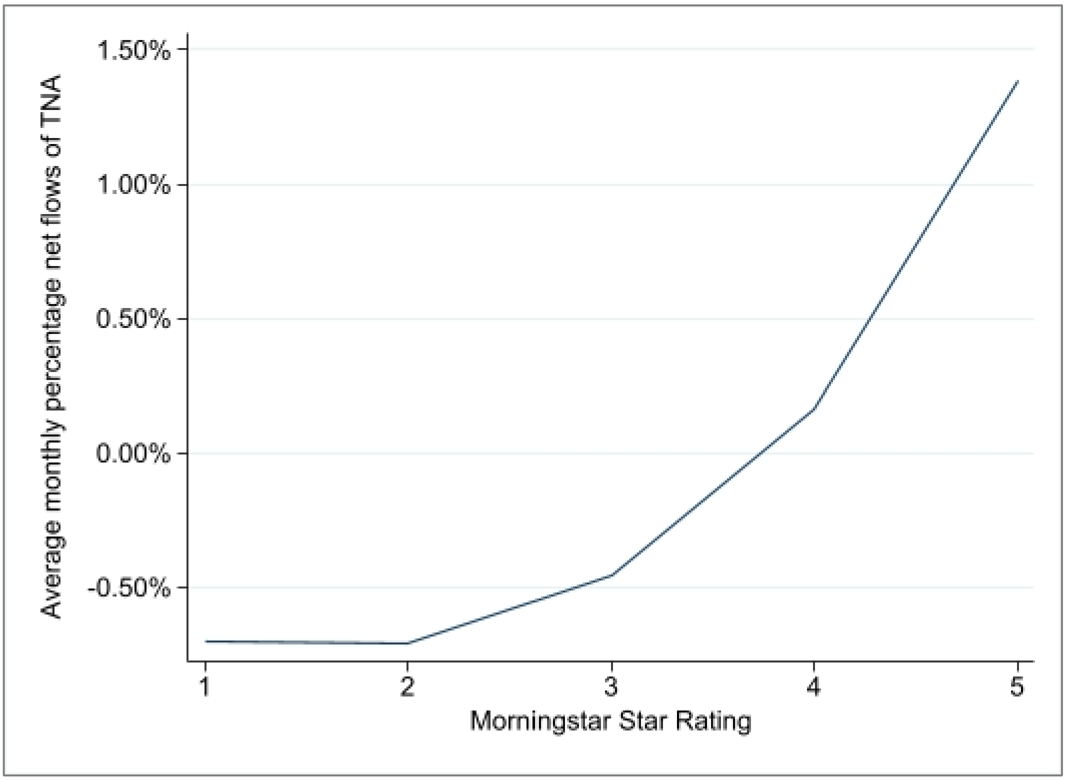
Accumulated returns on factor portfolio during the sample period

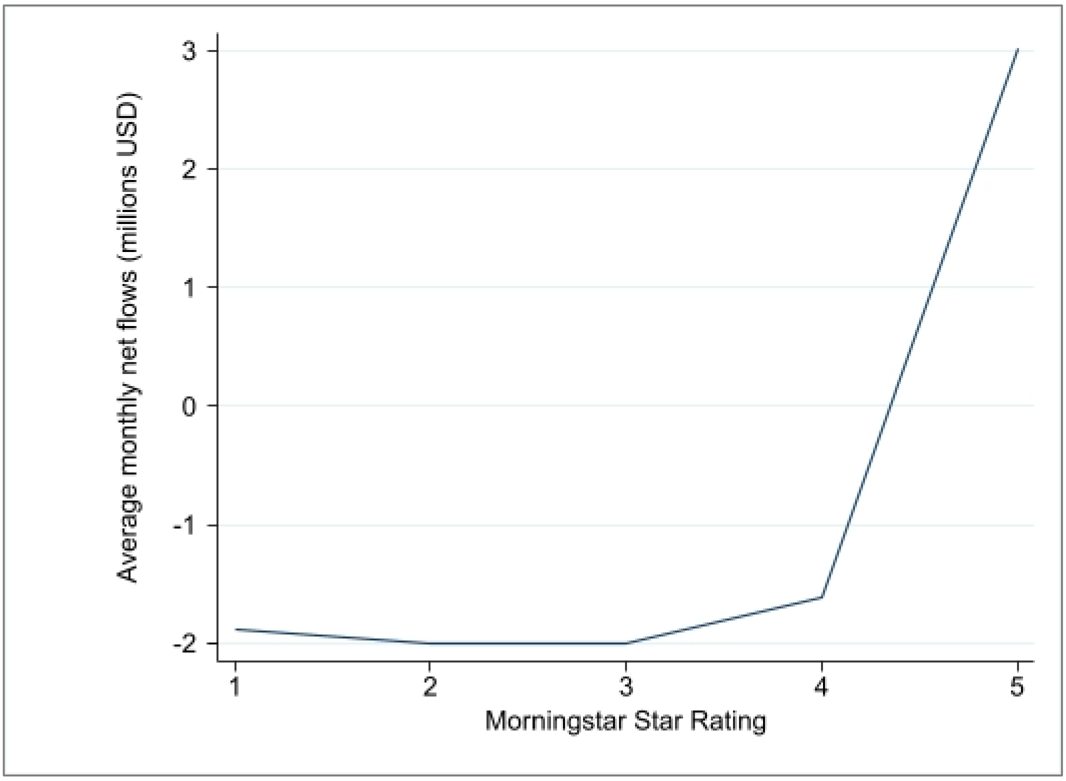


This figure presents the returns accumulated over the sample period for the factor portfolios for the developed market. All returns are indexed at 100 at the start of the sample period, the bold red line is a reference line at the index of 100. MktRf stands for the return on the market portfolio in excess of the risk-free rate, SMB stands for the return on the size factor portfolio (small minus big size), HML stands for the return on the value factor portfolio (high- minus low book-to-market ratio), WML stands for the return on the momentum factor portfolio (winners minus losers), and Rf stands for the risk-free rate.

Appendix B

Average percentage and dollar flows to Morningstar rated funds





These figures present the average percentage- and dollar flows to different star rated funds in the sample. These results are based on fund-month observations.

Appendix C

Robustness test of the decay rate on the flow-return component regression; model 2,3

	Model 2				
	$\hat{\lambda}$	$\hat{\lambda}$ *50%	$\hat{\lambda}$ *75%	$\hat{\lambda}$ *125%	$\hat{\lambda}$ *150%
	(1)	(2)	(3)	(4)	(5)
ALPHA	0.781***	0.849***	0.824***	0.732***	0.685***
	(0.048)	0.0563	0.0521	0.0443	0.0413
MKTRET	-0.030	-0.0173	-0.0242	-0.0334	-0.0358
	(0.050)	0.0438	0.047	0.0519	0.0532
SIZRET	0.766***	0.837***	0.806***	0.725***	0.687***
	(0.144)	0.1643	0.1539	0.1351	0.1276
VALRET	0.909***	0.971***	0.947***	0.865***	0.820***
	(0.112)	0.1255	0.1181	0.1067	0.1023
MOMRET	0.561***	0.564***	0.568***	0.546***	0.528***
	(0.130)	0.1506	0.1409	0.119	0.1097
Month fixed effects	No	No	No	No	No
Month-style fixed effects	Yes	Yes	Yes	Yes	Yes
Month-style-rating FE	No	No	No	No	No
Controls	Yes	Yes	Yes	Yes	Yes
Observations	207,189	207,189	207,189	207,189	207,189
R2	0.021	0.020	0.021	0.021	0.021
Adj. R2	0.012	0.012	0.012	0.012	0.012

	Model 3				
	$\hat{\lambda}$	$\hat{\lambda}$ *50%	$\hat{\lambda}$ *75%	$\hat{\lambda}$ *125%	$\hat{\lambda}$ *150%
	(1)	(2)	(3)	(4)	(5)
ALPHA	0.741***	0.797***	0.779***	0.697***	0.654***
	(0.048)	0.0557	0.0517	0.0442	0.0413
MKTRET	-0.021	-0.0073	-0.0146	-0.0247	-0.0275
	(0.052)	0.0455	0.0489	0.0542	0.0558
SIZRET	0.720***	0.782***	0.755***	0.682***	0.648***
	(0.146)	0.1674	0.1564	0.1369	0.1293
VALRET	0.868***	0.930***	0.907***	0.825***	0.781***
	(0.113)	0.127	0.1195	0.1079	0.1035
MOMRET	0.534***	0.520***	0.534***	0.524***	0.510***
	(0.135)	0.1541	0.1456	0.1248	0.1159
Month fixed effects	No	No	No	No	No
Month-style fixed effects	No	No	No	No	No
Month-style-rating FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
Observations	206,197	206,197	206,197	206,197	206,197
R2	0.047	0.046	0.047	0.047	0.047
Adj. R2	0.012	0.011	0.012	0.012	0.012

Appendix C (continued)

Robustness test of the decay rate on the flow-return component regression; model 2,3

This table presents robustness test of the panel regression results of fund flows on alpha and factor-related return components from the four-factor model, using different decay rates to estimate the independent variables. The model presented here are the second model, with month-style fixed effect; and the third model, with month-style-rating fixed effects. (1) presents the regression results when the decay rate $\hat{\lambda}$ = 0.20551497, the decay rate used in the main analysis. Columns (2), (3), (4) and (5) use $\hat{\lambda}$ times 50%, 75%, 125% and 150%, respectively. The control variables are the same as specified in Equation (8). Standard errors are double clustered by fund and month and are in parenthesis. *, ** and *** denote significance at the 10%, 5% and 1% level, respectively.