# Essays on Actively Managed Equity Mutual Funds

Andreas Ørpetveit

Dissertation submitted to the Norwegian School of Economics, for the degree of Philosophiae Doctor (PhD)

August 2021

# Contents

## Acknowledgements

First of all, I would like to express my deepest gratitude to my main supervisor Trond Døskeland for his support and guidance through my PhD. You have been available and encouraging from the beginning to the end of the PhD programme. I would also like to thank Petter Bjerksund for collaboration on one of the papers in my thesis, valuable feedback, and interesting discussions, as well as Jens Sørlie Kværner for helpful input on the papers in my thesis, and especially on my single-author paper. Collaborating with my office-mate André Wattø Sjuve has been motivating, and our common research interests have led to numerous interesting discussions and potential future research projects. I have been fortunate to become a part of a well-composed research group during my PhD, and I hope that many more papers will follow in the future.

Furthermore, I would like to thank all of my colleagues at the Department of Business and Management Science for a great and inspiring work environment. I am also grateful to the Norwegian School of Economics for financial support for this project.

Last but definitely not least, I want to express my great appreciation to the people closest to me. Especially, I would like to thank my fiancée, Benedicte, for putting up with me when I have been stressed. You have always supported me, and have made sure that I have kept a healthy work-life balance. I will also thank my parents, Bente and Per, for always supporting me and providing a safe place to come "home". Thank you for always being there for me!

Bergen, August 2021
Andreas Ørpetveit

# Introduction

This thesis focuses on actively managed equity mutual funds. In active funds, the fund managers make portfolio decisions on behalf of their investors. Their returns are usually evaluated against a benchmark index, where the benchmark is typically a market index. Thus, the funds sell the potential to outperform their benchmark, and investors pay a premium relative to the price of index funds for this service.[1] Fund managers may justify their fees if they can provide a realistic opportunity to outperform their benchmark. However, multiple studies have documented that active funds, on average, struggle to outperform their benchmark index after fees (see, e.g, Fama and French (2010), Ferreira, Keswani, Miguel and Ramos (2013) and Leippold and Rueegg (2020)).

During the past two decades there has been a shift in flows from actively managed funds toward passively managed index funds and ETFs (see, e.g., Cremers, Ferreira, Matos and Starks (2016)). The questionable track record of actively managed funds, which implies that the average investor has been better off by investing passively, might be one of the reasons for this shift. However, active mutual funds still manage trillions of dollars (Investment Company Institute (2020)). Moreover, the current trend is that people increasingly become their own money managers. At the same time, the financial landscape is becoming more complicated and many investors lack the necessary knowledge to judge the quality of the funds.[2] Bridging the information gap between fund managers and investors is therefore important to ensure investor protection. On the other hand, issues such as conflicts of interest, information asymmetry, and lack of financial sophistication among investors can increase the information gap and cause frictions.

Conflicts of interest in the mutual fund industry arise when the goals of fund managers and their investors are not aligned. Most mutual funds are run by a fund company (hereby referred to as a fund family). This family structure can lead to conflicts of interest, since the objective of the mutual fund family is to maximize the total revenue from its funds, while its investors seek to maximize the risk-adjusted returns on their investments (Chevalier and Ellison (1997)). This essentially means that the funds stand between serving their family with inflows and their investors with returns. Financial regulators have argued that sufficient level of industry competition may increase incentives for fund families to satisfy their clients' demands (see Australian Securities and Investment Commission (2020)).[3] If the quality of the active management service is affected by conflicts of interest, investors might end up paying for a dedicated asset management service that they do not receive.

Information asymmetry can increase incentives for opportunistic behavior by service providers. The existence of closet indexing provides a real-life example of such behavior. Closet indexers are funds with a low degree of active management, but are sold and marketed as actively managed funds (see, e.g., Cremers and Petajisto (2009) for US closet index funds and Cremers et al. (2016) for international closet index funds). The implication of closet indexing is that investors pay for an active portfolio management service, but receive a service that is close to an index fund. In the last couple of years,

[1] See, for example, Morningstar (2019).

[2] While also passive funds are a part of the financial landscape, investing in passive funds does not require the same evaluation of quality since the primary objective of index funds is to replicate the returns of a market index.

[3] The Australian Securities and Investment Commission (ASIC) commissioned a review of the level of competition in the Australian mutual fund industry (see ASIC) after the Financial Conduct Authority (FCA) in the UK found that price competition is weak in several areas of the UK fund industry (Financial Conduct Authority (2017)).

the issue of closet indexing has become a focus for financial supervisory authorities (FSAs) around the world, and the first FSAs to intervene toward closet index funds were in the Scandinavian countries.[4] In Norway, the intervention ended in a law-suit against one of the largest domestic funds, where the fund was convicted in the Norwegian Supreme Court and ordered to pay back 0.8% of the annual management fee from the period 2010 to 2014.[5]

The level of investors' financial sophistication affects to what extent their decisions are influenced by the information gap. The higher the level of financial sophistication among investors, the less likely they are to be affected by market frictions. Thus, investors can reduce the information gap themselves by improving their financial knowledge. Recently there has been a discussion on whether investors use a full asset pricing model (see Barber, Huang and Odean (2016) and Berk and van Binsbergen (2016)) or react to easily available signals (see, for example, Ben-David, Li, Rossi and Song (2019) and Evans and Sun (2021) for evidence on how Morningstar ratings drive fund flows) when deciding which funds to buy and sell.[6] This implies that if investors are able to interpret and use relevant information in the investment process, they are more likely to be able to make rational investment decisions.

The thesis is organized into three self-contained empirical papers, and has a particular focus on issues in the active mutual fund industry related to conflicts of interest from the family structure of the industry, closet indexing, and investors' financial sophistication. Below, I briefly describe the papers in the thesis.

In the first paper, I examine how fund families of actively managed funds respond to competition in terms of product development in an international fund sample. Fund family product development is defined as improving the quality of existing funds (e.g., level of activity, quality consistency, star funds, and manager changes) or as changes in the fund base (e.g., starting new funds, mergers, and liquidating funds). Thus, the quality channel entails making efforts to improve the alpha production in the family's funds, while the base channel is defined as expansions of the family's fund base. I find that when competition increases, fund families respond by increasing the quality of their existing funds rather than focusing on expanding their fund base. Furthermore, product development through the quality channel increases the performance of the family's funds, while product development through the base channel increases the flows to funds in the family as well as the family's market share. The last two results imply that, when fund families face greater competitive pressure, they choose to increase the quality and performance of their funds, and do not focus on increasing their market share. This response is in favor of the investors, and I therefore argue that competition reduces conflicts of interest that stem from the family structure of the industry.

The second paper is co-authored with Petter Bjerksund, Trond Døskeland, and André Wattø Sjuve. We examine the impact of policy scrutiny on Scandinavian closet index funds in a quasi-natural experiment. Closet indexers are defined as funds with active share (Cremers and Petajisto (2009)) below 40% and 50%.[7] In our experiment design, the treated funds are Scandinavian closet indexers, while

[4]See Financial Supervisory of Denmark (2013), Norwegian Ministry of Finance (2015), and Financial Supervisory of Sweden (2015) for details about the interventions in Denmark, Norway, and Sweden, respectively.

[5]This law-suit was the first of its kind worldwide. For more information about the verdict see Lovdata HR-2020-475-A.

[6]Morningstar ratings are purely quantitative, and the main input is the past performance. Thus, the results showing that investors respond to ratings might also be a manifestation of investors chasing past performance (see, for example, Sirri and Tufano (1998) and Ferreira, Keswani, Miguel and Ramos (2012)).

[7]Both cutoffs are used separately in the analyses throughout the paper. An active share of 50% essentially means that the fund invests 50% of its portfolio in the benchmark index and 50% of its portfolio in an active portfolio.

the control group are similar closet index funds from other European countries with no intervention from FSAs. The main finding is that funds under scrutiny chose to increase active share rather than to reduce fees and update their investor information. Furthermore, we find that the value creation in the funds under scrutiny decreased after they increased activity. Therefore, the investors would be better off if the funds responded by reducing fees and updating investor information. The reduction in performance suggests that the fund managers chose to follow a closet indexing strategy because they lack sufficient skill or investment ideas to have a higher active share. Based on our findings, we argue that regulators should motivate closet index funds to reduce fees rather than to increase activity.

Finally, the third paper is co-authored with Trond Døskeland and André Wattø Sjuve. We develop a model showing that the level of active management and fund fee are valuable signals for a fund's potential to beat its benchmark index. Active fee (Cremers and Quinn (2016)) is constructed as a combination of active share and total expense ratio, and can be interpreted as the unit price of active management.[8] This signal is somewhere "in the middle" of the full asset pricing models and the easily accessible signals, such as Morningstar ratings. We find a negative time-series relationship between active fee and subsequent net flows, which can be interpreted as a negative price elasticity of demand for active management. These results are driven by both a positive active share-flow relationship and a negative fee-flow relationship. Moreover, while Morningstar ratings have a high standing in the mutual fund industry, our results also hold when controlling for these in our regressions. Our results suggest that investors are rational in the sense that they are able to interpret and use information about active share and fee when buying and selling funds.

[8] In our model, the lower active fee, the larger is the upside potential of the fund.

## References


Australian Securities and Investment Commission (2020), 'ASIC MARKET STUDY CONSULTANCY FUNDS MANAGEMENT INDUSTRY TERMS OF REFERENCE '. Available at https://download.asic.gov.au/media/5727060/terms-of-reference-review-of-competition-in-the-australian-funds-management-industry.pdf, Last accessed on 2021-16-06.

Barber, B. M., Huang, X. and Odean, T. (2016), 'Which Factors Matter to Investors? Evidence from Mutual Fund Flows', *The Review of Financial Studies* **29**(10), 2600–2642.

Ben-David, I., Li, J., Rossi, A. and Song, Y. (2019), 'What do mutual fund investors really care about?', *Fisher College of Business Working Paper* **118**(2019-03), 005.

Berk, J. B. and van Binsbergen, J. H. (2016), 'Assessing asset pricing models using revealed preference', *Journal of Financial Economics* **119**(1), 1–23.

Chevalier, J. and Ellison, G. (1997), 'Risk taking by mutual funds as a response to incentives', *Journal of political economy* **105**(6), 1167–1200.

Cremers, K. M. and Petajisto, A. (2009), 'How active is your fund manager? a new measure that predicts performance', *Review of Financial Studies* **22**(9), 3329–3365.

Cremers, K. and Quinn, C. (2016), 'Do mutual fund investors get what they pay for: Securities law and closet index funds', *Virginia Law and Business Review* **11**, 31.

Cremers, M., Ferreira, M. A., Matos, P. and Starks, L. (2016), 'Indexing and active fund management: International evidence', *Journal of Financial Economics* **120**(3), 539–560.

Evans, R. B. and Sun, Y. (2021), 'Models or Stars: The Role of Asset Pricing Models and Heuristics in Investor Risk Adjustment', *The Review of Financial Studies* **34**(1), 67–107.

Fama, E. F. and French, K. R. (2010), 'Luck versus skill in the cross-section of mutual fund returns', *Journal of Finance* **65**(5), 1915–1947.

Ferreira, M. A., Keswani, A., Miguel, A. F. and Ramos, S. B. (2012), 'The flow-performance relationship around the world', *Journal of Banking & Finance* **36**(6), 1759–1780.

Ferreira, M. A., Keswani, A., Miguel, A. F. and Ramos, S. B. (2013), 'The determinants of mutual fund performance: A cross-country study', *Review of Finance* **17**(2), 483–525.

Financial Conduct Authority (2017), 'Asset Management Market Study'. Available at https://www.fca.org.uk/publication/market-studies/ms15-2-3.pdf, Last accessed on 2021-16-06.

Financial Supervisory of Denmark (2013), 'Market developments 2013'. Available at https://www.dfsa.dk/~/media/Tal-og-fakta/2014/Markedsudvikling/Market-developments-2013-Investment-associations.pdf?la=en, Last accessed on 2021-23-07.

Financial Supervisory of Sweden (2015), 'Consumer protection on the financial market'. Available at https://www.fi.se/contentassets/9791b37d1d074ed4a3fc193786d449f6/konsumentrapp_2015engny.pdf, Last accessed on 2021-23-07.

Investment Company Institute (2020), 'A Review of Trends and Activities in the Investment Company Industry'. Available at https://www.ici.org/system/files/attachments/pdf/2020_factbook.pdf, Last accessed on 2021-23-07.

Leippold, M. and Rueegg, R. (2020), 'How rational and competitive is the market for mutual funds?', *Review of Finance* **24**(3), 579–613.

Morningstar (2019), 'U.S. Fund Fee Study'. Available at https://www.morningstar.com/articles/925303/2018-morningstar-fee-study-finds-that-fund-prices-continue-to-decline, Last accessed on 2021-23-07.

Norwegian Ministry of Finance (2015), 'Financial Markets Report 2015'. Available at https://www.regjeringen.no/contentassets/fcbb121a9cb342f0b52d44300d3cc3b0/no/pdfs/stm201520160029000dddpdfs.pdf, Last accessed on 2021-23-07.

Sirri, E. R. and Tufano, P. (1998), 'Costly search and mutual fund flows', *The journal of finance* **53**(5), 1589–1622.

# Competition and Fund Family Product Development

Andreas Ørpetveit*
Norwegian School of Economics

**Abstract**

Despite extensive evidence of how mutual fund competition affects fund fee and performance outcomes, there is little evidence of how competition affects the incentives of market participants. This paper uses an international sample of active equity mutual funds to examine how product development in mutual fund families is affected by competitive pressure. Fund family product development is defined as improving the quality of existing funds (e.g., level of activity, quality consistency, star funds, and manager changes) or as changes in the fund base (e.g., starting new funds, mergers, and liquidating funds). The results show that greater industry competition motivates fund families to carry out product development through the quality channel rather than the base channel. Furthermore, product quality development increases performance in the family-affiliated funds, and thus benefits the investors. Based on the findings, I argue that competition motivates desired activity in the mutual fund industry and reduces conflicts of interest that stem from the family structure of the industry.

**JEL Classification:** G11; G20; L10

**Keywords:** Mutual funds; mutual fund families; competition; product development

*Email: andreas.orpetveit@nhh.no.

# 1 Introduction

The implications of competition in the active mutual fund industry have been the subject of multiple academic studies. However, these studies primarily address fund fee and performance outcomes as a result of competition.[1] To understand how competition affects industry efficiency, one must also understand how competition affects the incentives of market participants, i.e., whether a competitive industry drives desired activity or distorts incentives. Mutual funds manage trillions of dollars. An efficient mutual fund industry is therefore important to ensure investor protection and financial stability. More specifically, for competition to be a potential source of industry efficiency, it should motivate the market participants to make efforts that benefit investors. In contrast, Feldman, Saxena and Xu (2020) find that fund managers reduce their active share, i.e., their alpha production efforts, when competition increases. Their findings suggest that competition might distort incentives in a way that does not benefit investors.

This paper contributes to the literature on the implications of competition for incentives by examining how mutual fund companies respond to competition in terms of product development. I use an international sample of equity funds from 40 countries, and focus exclusively on actively managed funds, since passive index funds are more homogeneous products without the objective to outperform a benchmark. Moreover, I focus on mutual fund companies (hereby referred to as mutual fund families) as opposed to individual funds because it is at the family-level the strategic decision making takes place (Khorana and Servaes (2012)).[2]

The family structure in the mutual fund industry can lead to conflicts of interest. While the objective of the mutual fund family is to maximize the total revenue from its funds, i.e., assets under management multiplied by fees, its investors seek to maximize the risk-adjusted returns on their investments (Chevalier and Ellison (1997)). In light of these potential conflicts of interest, I examine whether greater competitive pressure distorts or motivates desired activity in fund families. The results are also relevant to the literature on the role of the fund family in the mutual fund industry, i.e., to what extent fund family product development strategies affect the funds in the family.

Competition requires that firms have incentives to compete (Schmalz (2017) and Azar, Schmalz and Tecu (2018)). The level of mutual fund industry competition varies around the world, and industry characteristics such as common law, stock market turnover, quality of institutions, and regulation affect the level of competition (Ferreira and Ramos (2009)). The US mutual fund industry is regarded as one of the most competitive in the world, but researchers have debated whether the US mutual fund industry is competitive as a whole.[3] Separately from the academic literature, financial regulators have carried out investigations of the level of competition in the industries for which they are responsible. The Financial Conduct Authority (FCA) in the UK found that price competition is weak in several areas of the UK fund industry (Financial Conduct Authority (2017)). The FCA's findings from the UK prompted the Australian Securities and Investment Commission (ASIC) to commission a similar review

[1] See, e.g., Coates IV and Hubbard (2007), Gil-Bazo and Ruiz-Verdú (2009), Khorana, Servaes and Tufano (2009), Wahal and Wang (2011), and Khorana and Servaes (2012) for studies of fees and Dyck, Lins and Pomorski (2013), Pástor, Stambaugh and Taylor (2015), Hoberg, Kumar and Prabhala (2018), and Leippold and Rueegg (2020) for performance studies.

[2] In addition, more than 90% of the funds in the sample used in this paper are member of a fund family.

[3] See, for example, Coates IV and Hubbard (2007), Gil-Bazo and Ruiz-Verdú (2009), and Wahal and Wang (2011).

of their mutual fund industry.[4] ASIC underlines the importance of efficient competition in its terms of reference (Australian Securities and Investment Commission (2020)): "Effective competition implies that firms have sufficient incentives to identify and satisfy clients' demands as efficiently as possible and constantly seek to win the business of clients who use rival services." Based on these examples, I argue that it is important to identify the channels through which competition benefits investors if we are to understand how to stimulate the industry when the level of competition is sub-optimal.

To test how fund families adjust product development to deal with competition, I begin by constructing product development variables. Defining product development can be difficult. In general, it is the term for creating an improved or new product and delivering it to the market (Marxt and Hacklin (2005)).[5] In fund families, improving products entails making efforts to improve the alpha production or risk-adjusted returns in their funds. This is defined as a product quality development channel. Creating new products expands the product base, and can be implemented by starting new funds or expanding the fund base into new categories. This is defined as a product base development channel. Since product base development is defined as an expansion of the product base, starting a new fund contributes positively, while a fund liquidation contributes negatively. Looking to the academic literature on asset management, I identify four fund family variables within the product quality channel and five fund family variables within the product base channel.[6] To structure the analysis, I construct one family-level index for each channel. First, for each of the nine variables, the family is ranked in the entire cross-section each year. Next, the quality (base) index is constructed as the equal-weighted average of the family-level quality (base) input variable ranks.

Two industry competition variables are defined, both based on standard industry concentration measures. The first is the Herfindahl-Hirschman index (*HHI*) subtracted from 1, while the second corrects for the number of firms and is the normalized Herfindahl-Hirschman index (*NHHI*) subtracted from 1. Using the product development indices and industry competition variables, I test whether fund families adjust their product development in response to competition. When competition increases, the fund families face greater competitive pressure from rival families and have to differentiate in order to stay competitive. While product development through the quality channel consists of efforts to differentiate in terms of quality, product development through the base channel consists of efforts to differentiate in terms of the products offered by the family. I hypothesize that fund families increase product development through both channels when competition increases.

The hypothesis is tested using two different empirical methodologies. First, I run family fixed effects panel regressions to test how the within-family product development changes in response to competition. The results show that greater competition is associated with an increase in product quality development. A one standard deviation increase in industry competition corresponds to a 27% increase in the product quality development index, compared to the sample average. I also find weak statistical evidence for product base development declining when competition increases. The results are similar for both of the industry competition variables. Furthermore, the results hold in

[4] More information on the upcoming report from the Australian mutual fund industry can be found at ASIC.

[5] See also Ullman (1992) and Ulrich (2003) for discussions of the definition of product development.

[6] The four product quality development variables are the activity level of the funds in the family, fund manager turnover, family performance dispersion, and family star fund creation. The five product base development variables are fund starts, fund liquidations, within-family mergers, fund category starts, and fund category liquidations. All variables are constructed at the fund family-level.

cross-sectional first-difference tests, at both the industry- and family-level.

Second, to address potential endogeneity issues in the industry competition measures, I follow the methodology of Rodríguez-Castelán, López-Calva and Barriga-Cabanillas (2019) and treat the industry competition variables as endogenous and instrument them using Bartik-type instruments to perform instrumental variable (IV) regressions (see Bartik (1991) and Bartik (2002)). This approach isolates the exogenous source of variation in industry competition. The growth in *HHI* and *NHHI* is instrumented by the growth in the US active mutual fund industry and the global active mutual fund industry, in two separate IV-procedures. The direction of the IV-results for the product quality development index corresponds with the main results, and the results are around twice as strong in terms of magnitude and economic significance. Furthermore, the coefficients of competition on product base development change signs, but are not statistically significant in these regressions. Based on these results, I argue that the true effect of industry competition on product quality development is larger than estimated in the main tests, and that industry competition does not appear to affect product base development.

These results also highlight the importance of accounting for endogeneity when testing how industry competition affects firm-level variables in order to avoid biased estimates. The relationship between industry competition and product quality development is an interesting finding because previous studies have found that fund families tend to expand their product base as a growth strategy (see e.g., Khorana and Servaes (1999) and Khorana and Servaes (2012)). In contrast, these results suggest that, to stay competitive, fund families prefer to increase the overall quality of their funds.

After establishing how competition affects fund family product development, I test the outcomes of product development on value creation in family-affiliated funds and the revenue of the family. Based on the nature of the input variables in the product development indices, I hypothesize that product development through the two channels improve the family's competitiveness in different ways. First, I hypothesize that product quality development increases the performance of family-affiliated funds. Second, I hypothesize that product base development increases the market share of the family, i.e., makes the fund family more competitive in terms of industry position and revenue. These tests essentially test the fund families' motives for carrying out product development in the two channels.

The results support the hypotheses. I find that product quality development improves the performance of the funds in the family, and that product base development increases the flows to family-affiliated funds and the family's market share. While I find no evidence that product base development enhances the performance of family-affiliated funds, I find some evidence that product quality development increases flows to the family's funds. This might be attributed to the fact that investors tend to chase past performance when selecting funds (see, e.g., Sirri and Tufano (1998) and Ferreira, Keswani, Miguel and Ramos (2012)). These results imply that, when fund families face increasing competitive pressure, they choose to increase the quality and performance of their funds, and do not focus on increasing their market share. Product development responses to competition are therefore in favor of the investors, and I argue that this is evidence that competition reduces conflicts of interest from the family structure in the mutual fund industry. While Feldman et al. (2020) find that fund managers are less willing to make alpha production efforts when competition increases, I do not find evidence for this in my family-level analysis. However, this analysis has two key differences. First, the quality development index does not include portfolio-specific variables, and, second, I focus on fund family-level efforts and not on efforts in individual funds. Thus, I argue that these results are not in direct contrast to their findings.

This paper makes two contributions. First, it adds to the literature on how market participants in the mutual fund industry are affected by competition. As mentioned in the introduction, previous studies of competition in the mutual fund industry have mainly focused on fee and performance outcomes. Studies from outside the mutual fund literature find that competition can affect or distort incentives.[7] The results in this paper suggest that competition is a driver of optimal behavior in the mutual fund industry, as it improves investor welfare. Khorana and Servaes (2012) find that fund families compete in a non-price dimension, mainly through a product differentiation channel. Cremers, Ferreira, Matos and Starks (2016) and Sun (2020) study the effect of competition from passively managed funds and find that active funds differentiate at the portfolio-level in response to growth in passive alternatives. My results show that the fund families differentiate in terms of quality when competition increases.

Second, the results also shed light on the role of the mutual fund family and how the family structure of the industry affects the funds. Since fund families are responsible for strategic decisions, such as overall family product development, it is important to know how this structure affects the funds. My findings show that fund performance increases with product quality development, and underline the role of the fund family in fund performance found in previous studies. For example, Fang, Kempf and Trapp (2014) and Berk, Van Binsbergen and Liu (2017) find that fund families allocate their fund managers strategically to maximize the total value creation in the family. Other studies find a significant impact of the fund family on performance and fund portfolios (see, e.g., Guedj and Papastaikoudi (2003), Gaspar, Massa and Matos (2006), Elton, Gruber and Green (2007)), Kempf and Ruenzi (2008), Pollet and Wilson (2008), and Chan, Lai and Lee (2017)). However, to my knowledge, this is the first paper to connect the overall fund family product development strategy with performance.

Massa (2003) argues that, the more families can differentiate themselves in terms of non-performance related characteristics, the less they need to compete in terms of performance. My results show that competition motivates fund families to carry out product quality development. The quality channel of product development increases the performance of family-affiliated funds and makes the fund families more competitive in terms of performance. Product base development increases inflows to family-affiliated funds and the market share of the family, but this is not the main focus of fund families when the competitive pressure increases. I argue that the results are evidence of competition being a way of reducing the conflicts of interest that stem from the family structure of the industry.

The rest of the paper is structured as follows. Section 2 develops the hypotheses. Section 3 presents the data, summary statistics, and the empirical method. In Section 4, I test how fund families respond to competition. In Section 5, I test how product development affects fund performance and fund family market share. Section 6 discusses the results and concludes.

[7] See, for example, Karuna (2007), Becker and Milbourn (2011), Bennett, Pierce, Snyder and Toffel (2013), Cornaggia, Mao, Tian and Wolfe (2015), Chhaochharia, Grinstein, Grullon and Michaely (2017), Schmidt, Fey and Thoma (2017), Aghion, Bechtold, Cassar and Herz (2018), and Bustamante and Frésard (2020) for studies from outside the mutual fund literature on how competition can affect incentives.

# 2 Hypothesis development

In this section, I define the industry competition variables, describe the motivation for and derive the product development indices, and develop the testable hypotheses.

## 2.1 Industry competition

Industries are defined as the funds' countries of domicile.[8] The two industry competition variables are based on the industry Herfindahl-Hirschman index (*HHI*) and the normalized Herfindahl-Hirschman index (*NHHI*). They differ in the sense that *NHHI* adjusts for the number of firms in the industry, while the number of firms can influence *HHI*. More specifically, the number of firms affects the possible range of *HHI*, which implies that industries with many firms are assigned a lower index value based on the number of firms, and not necessarily the level of competition itself. The *HHI* variable of industry concentration is well-grounded in industrial organization theory (see Tirole and Jean (1988)), and is defined as the sum of the squared market weights of fund families within the industry. The *HHI* industry concentration in industry $c$ at time $t$ is given by

$$HHI_{c,t} = \sum_{f} w_{f,c,t}^2, \tag{1}$$

where $w_{f,c,t}$ is the weight of family $f$ in industry $c$ at time $t$. *NHHI* is defined by

$$NHHI_{c,t} = \frac{HHI_{c,t} - \frac{1}{N_{c,t}}}{1 - \frac{1}{N_{c,t}}}, \tag{2}$$

where $N_{c,t}$ is the number of firms in industry $c$ at time $t$. By definition, a reduction in *HHI* and *NHHI* corresponds to an increase in competition. Thus, for interpretation purposes, the competition variables are defined by subtracting the industry concentration from 1.[9] Then, the competition variable based on *HHI*, $Comp^{HHI}$, is defined by

$$Comp_{c,t}^{HHI} = 1 - HHI_{c,t}. \tag{3}$$

The competition variable based on *NHHI*, $Comp^{NHHI}$, is defined by

$$Comp_{c,t}^{NHHI} = 1 - NHHI_{c,t}. \tag{4}$$

## 2.2 Product development indices

In this section, I define the product development indices. The definition of product development is to improve existing products or to develop new products (see, e.g., Ullman (1992), Ulrich (2003), and Marxt and Hacklin (2005) for definitions of product development). Improvement of existing products is defined as a product quality development channel, while the development of new products is defined as a product base development channel. In fund families, the quality channel includes efforts made to

[8] Countries of domicile represent the country in which a fund is legally organized and are often used to define home countries in international mutual fund studies (see, for example, Cremers (2016) and Demirci, Ferreira, Matos and Sialm (2020)).

[9] This adjustment is often done to make the Herfindahl index coincide with the level of competitive pressure (see, for example, Moshirian, Tian, Zhang and Zhang (2021)).

increase the quality of existing funds, while the base channel includes expansions of the family's fund base. I construct one index for each channel, consisting of the decile rank of relevant input variables.[10] In addition, I use these two indices to construct a combined product development index and a net quality development index. Appendix A provides a detailed example of how the indices are constructed for one of the families in the sample, as well as summary statistics for the product development indices.

### 2.2.1 Product quality development

First, I describe the motivation for and define the input variables in the product quality development index. The input variables include family-level variables that can describe the quality, consistency of quality, or efforts to increase the quality of the funds in the family. Looking to the mutual fund literature, I identify four variables that I construct at the family-level: level of active management, fund manager turnover, within-family performance dispersion, and share of star funds. The index is constructed annually for each fund family by computing decile ranks of the input variables in the entire cross-section each year. Variables defined as the share of funds are divided by the number of funds in the family at the end of the previous year. Below, I describe the motivation for the input variables.

Multiple studies find that the level of active management can predict fund performance (see Wermers (2003) for tracking error, Cremers and Petajisto (2009) for active share, and Amihud and Goyenko (2013) for $R^2$). I define the activity level of a fund by its tracking error, i.e., the standard deviation of the benchmark-adjusted returns. To compute a family's decile rank, I first compute each fund's decile rank and average this across the funds in the family.

Changing fund managers can be a way for a fund family to improve its funds. Bessler, Blake, Lückoff and Tonks (2018) find that funds with low past performance experience improved subsequent performance after manager replacements. A number of studies document an inverse relationship between past fund performance and manager changes (see, e.g., Khorana (1996), Chevalier and Ellison (1999), and Kostovetsky (2017)). Bryant (2012) argues that conflicts of interest affect the decisions to replace fund managers, because high expense ratio fund managers have a lower probability of replacement for a given level of underperformance. I define the manager turnover variable as the share of funds with manager changes during a year.[11] This variable measures how willing a fund family is to reallocate its fund managers and take action against underperformance.

Fund families might have many funds in order to increase the probability of increasing assets under management, and not necessarily to run them optimally. To account for this, I include a variable of within-family performance dispersion. I define this by the standard deviation of the cross-sectional benchmark-adjusted returns within the family each year. Using the benchmark-adjusted returns adjusts for the differences in "segments" or categories, such that it measures the quality consistency across funds in the family. Massa (2003) finds that performance dispersion at the family-level negatively affects performance. Based on this, I rank the families such that the families with the lowest (highest) dispersion, i.e., highest (lowest) quality consistency, are assigned the highest (lowest) decile ranks.

Next, I construct a family-level star fund measure. This is motivated by the findings of Khorana

[10] The construction of the indices is similar to the construction of the competition and cooperation indices of Evans, Prado and Zambrana (2020).

[11] If any funds in the family are missing manager data, the number of manager changes is divided by the number of funds with manager data.

and Servaes (2012), who find that the presence of a star fund has a positive and significant impact on the fund family market share. Nanda, Wang and Zheng (2004) find that the presence of star funds in a family attract flows to other member funds as well. I define star funds as the funds with past year benchmark-adjusted returns in the top 5th percentile within their domicile-category.[12] The star fund variable is then defined as the share of funds defined as a star fund.

Finally, with all variables defined, I construct the index using decile ranks of the input variables explained above. The product quality development index is computed as the equal-weighted decile rank over the input variables

$$Quality_{f,c,t} = \frac{1}{4}\left(TE^{DecRnk}_{f,c,t} + Manager^{DecRnk}_{f,c,t} + Dispersion^{DecRnk}_{f,c,t} + Star^{DecRnk}_{f,c,t}\right), \tag{5}$$

where $TE^{DecRnk}_{f,c,t}$ is the decile rank of the family tracking error, $Manager^{DecRnk}_{f,c,t}$ the decile rank of manager changes, $Dispersion^{DecRnk}_{f,c,t}$ is the decile rank of family performance dispersion, and $Star^{DecRnk}_{f,c,t}$ is the decile rank of the share of star funds. All input variables range from 0 to 10, where families with zero manager changes and zero star funds are given the value 0.[13] The nature of the index is such that it cannot be constructed for families with fewer than two funds. Families with missing data for any of the input variables are not given a decile rank for these input variables, and will therefore not be assigned a product quality development index.

### 2.2.2 Product base development

Next, the motivation for the product base development index is described. This index describes the family's fund base expansions. The input variables include variables for fund starts, fund liquidations, and within-family fund mergers. Since the index is defined as an expansion variable, starting a new fund counts positively, while a fund liquidation counts negatively toward the index. All input variables are defined as shares of the number of funds in the family at the end of the previous year.[14] The index is constructed in the same way as the product quality development index, i.e., by computing decile ranks of the input variables in the entire cross-section each year.

There are three reasons why families open new funds (Khorana and Servaes (2012)). First, a family may want to offer a new fund with variation from existing products. The second reason is to increase the likelihood of having a top-performing fund. Finally, families may open new funds because their current funds have performed poorly and fail to attract inflows. Khorana and Servaes (1999) find that the decision to start new funds is motivated by the potential to generate income, and that families with low fees and families with star funds are more likely to open new funds. I divide fund starts into two separate variables, where the first comprises funds started in fund categories where the family already has funds and the second comprises new funds in new fund categories.[15] For both variables, I define the variable as the share of funds at the end of the previous year.

Fund base reductions can take the form of mergers or fund liquidations. First, within-family mergers are defined. I focus exclusively on within-family mergers because they only involve the decision of

[12]This definition of star funds follows from Nanda et al. (2004).
[13]Families with missing manager history data for all funds are excluded and not given a decile rank.
[14]This is done to account for larger families having more changes in their fund base.
[15]These two variables do not overlap, meaning that a fund start is only placed in one of the groups, not both.

the family of the target fund. Prior research shows that fund mergers are often the result of poor performance (Jayaraman, Khorana and Nelling (2002) and Zhao (2005)). Since within-family mergers reduce the fund base, this variable counts negatively toward the product development index. The merger variable is therefore defined as the negative sum of within-family mergers divided by the number of funds in the family at the end of the previous year.

Next, I define the fund liquidation variables. As for fund starts, I define two separate variables: funds liquidated in fund categories where the family still has funds after the liquidations, and funds where the fund liquidation entails a category liquidation. Zhao (2005) studies exit decisions in the US mutual fund industry and finds that liquidated funds tend to be small and younger funds, and that a family tends to liquidate relatively unique portfolios to stay focused. Like the merger variable, I define the fund and category liquidation variables as the negative sums of liquidations divided by the number of family funds at the end of the previous year.

The product base development index is constructed using the decile ranks of the input variables explained above, and the index is defined as the equal-weighted decile rank over the input variable ranks

$$Base_{f,c,t} = \frac{1}{5}(FundStart_{f,c,t}^{DecRnk} + FundLiq_{f,c,t}^{DecRnk} + Mergers_{f,c,t}^{DecRnk} + CatStart_{f,c,t}^{DecRnk} + CatLiq_{f,c,t}^{DecRnk}), \quad (6)$$

where $FundStart_{f,c,t}^{DecRnk}$, $FundLiq_{f,c,t}^{DecRnk}$, $Mergers_{f,c,t}^{DecRnk}$, $CatStart_{f,c,t}^{DecRnk}$, and $CatLiq_{f,c,t}^{DecRnk}$ are the decile ranks of fund starts, fund liquidations, within-family mergers, category starts, and category liquidations, respectively. All input variables range from 0 to 10. Families with zero new funds or zero new categories are given decile rank of 0 for the variables equal to zero. Families with zero within-family mergers, fund liquidations, or category liquidations are given decile rank of 10 for the variables equal to zero.

For both the $Quality_{i,t}$ and $Base_{i,t}$ indices, I standardize the values annually across all families (with mean of zero and standard deviation of one) to isolate the cross-sectional differences in the indices across families. To distinguish between the two indices and make sure that they do not influence each other, all variables in the quality index only contain funds with complete data within a year, i.e., they do not include funds that were started or liquidated during a year.

Next, using the product quality and base development indices, I construct a total product development index and a net quality development index. These indices are used to examine how the two channels of product development are related to each other, and how they affect other outcome variables in combination. The total product development is the measure of overall product development for the family, and it takes both of the channels into account. It is defined as the sum of the quality and the base index

$$Development_{f,c,t} = Quality_{f,c,t} + Base_{f,c,t}. \quad (7)$$

The net quality index measures the product quality development net of the product base development and describes whether fund families use one of the channels to scale up the other channel. Moreover, in the regressions, this index is used to test the differences in how the two channels are affected by competition and how they affect fund family outcomes. The net quality index is defined as

$$Net\ Quality_{f,c,t} = Quality_{f,c,t} - Base_{f,c,t}. \tag{8}$$

Appendix A.1 provides a detailed example of how the indices are constructed for one of the families in the sample. Appendix A.2 presents summary statistics for the product development indices, with distribution of deviation from within-family means in Figure A1 and the equal-weighted and value-weighted time-series of average deviations from within-family means in Figure A2. Next, I derive the testable hypotheses based on the industry competition variables and the product development indices.

### 2.3 Hypotheses

Fund families are responsible for the strategic decisions in the mutual fund industry, and this family structure of the industry can lead to conflicts of interest. The source of these potential conflicts is that a mutual fund family's objective is to maximize the total revenue from its funds, while its investors want the fund family to maximize alpha production (Chevalier and Ellison (1997)). The underlying economics of the hypotheses concerns how fund family product development is affected by competition, and whether competition triggers or reduces conflicts of interest.

#### 2.3.1 Competition and product development

When competitive pressure in the industry increases, investors have more investment alternatives, and it becomes more difficult for the families to maintain their market share. Thus, in order to maintain their position in the industry or acquire additional market share, the families have to put more effort into differentiating themselves from rival families. I conjecture that the objective of differentiation is to stay competitive, and that both of the product development channels derived in the previous section are applicable family strategies for differentiation from rival families.

Fund family product quality development consists of efforts made to improve the quality or quality consistency of the funds in the family. While product development through this channel is not necessarily directly observable for investors, increased performance may well be the best way for the family to differentiate. The perception that investors follow past performance also has empirical backing (see, e.g., Sirri and Tufano (1998) and Ferreira et al. (2012)).[16] In addition, if some of the funds in the family perform well, this may have spillover effects in the form of flows to other funds in the family.[17]

Fund family product base development consists of efforts made to expand the family through starting new funds. The nature of this channel is such that families can expand into new categories or start new funds in categories where their funds are successful. In terms of making the family more competitive, product development through this channel increases the family's presence, because there will be more funds from the family for investors to choose from. Fund liquidations reduce the presence of the family, and therefore count negatively toward the product base channel of product development. Khorana and Servaes (2012) argue that fund families compete in a non-price dimension, and find that starting new funds is an effective way to differentiate the family and grow.

[16] A functioning flow-performance relationship is also an essential part of the theoretical model of Berk and Green (2004).
[17] See, for example, Nanda et al. (2004) for examples of such spillover effects in families with star funds.

Other studies examine how active mutual funds differentiate in response to increasing competition from passive investment alternatives. Cremers et al. (2016) and Sun (2020) find that active funds tend to differentiate themselves from passive funds by increasing their degree of active management. In other words, the active funds make efforts to differentiate their portfolios from the portfolios of passive funds. However, passive and active funds are not perfect substitutes, while the products in families of active funds are closer to being perfect substitutes for active funds from other families. Therefore, it may be more important for fund families to differentiate at the family-level. Building on these findings, I hypothesize that greater competitive pressure force the families to make efforts to differentiate, and that both the quality and base channels are applicable strategies for this. I formulate the first testable hypothesis in the following two sub-hypotheses:

**Hypothesis 1a: Fund families respond to increasing competitive pressure by increasing product quality development.**

**Hypothesis 1b: Fund families respond to increasing competitive pressure by increasing product base development.**

### 2.3.2 Product development outcomes

As mentioned in the motivation for the first hypothesis, product development strategies are ways for a the family to differentiate itself from rival families. However, there can be two main motives for product development. First, the family might want to become more competitive in terms of performance, and, second, they might want to become more competitive in terms of revenue and market position. While the first motive may benefit both the investors and the family itself, the latter motive entails making efforts to increase the total assets under management and does not necessarily benefit the investors.

This hypothesis concerns the potential conflicts of interest in the mutual fund industry, where a family and its investors can have conflicting motives. For competition to be a potential source of industry efficiency, fund family responses to competition should benefit investors and reduce conflicts of interest. If the fund families take measures in order to increase the income of the family and not the value for investors, competition reduces industry efficiency. Massa (2003) argues that, the more families can differentiate themselves in terms of nonperformance-related characteristics, the less they need to compete in terms of performance. Therefore, I conjecture that the two channels of product development are based on different motives.

The nature of the product quality development index is such that these are efforts made to improve the overall quality of the value creation and performance in the family's funds. I therefore conjecture that the main goal for the family when carrying out product quality development is to increase the performance of its funds. Previous studies have found that the fund family plays a significant role for the performance of its funds, both by affecting the risk taking in its funds (see, e.g., Elton et al. (2007), Kempf and Ruenzi (2008), and Chan et al. (2017)) and by making strategic decisions (see, e.g., Guedj and Papastaikoudi (2003), Gaspar et al. (2006), Fang et al. (2014), and Berk et al. (2017)). Therefore, the first part of this hypothesis concerns how the product quality development channel affects the performance of the funds in the family. I hypothesize that product development through the quality channel increases performance in the family's funds.

On the other hand, the product base development index does not include changes or improvements to existing funds, and it is separate from the quality channel in the sense that the measures do not overlap.[18] It is therefore not likely that this product development channel will directly affect the performance of the funds in the family. I conjecture that the main goal for the family when carrying out product base development is to increase the family's market share and revenue.

In a study of non-portfolio related differentiation in homogeneous S&P 500 index funds, Hortaçsu and Syverson (2004) find that investors value funds' observable non-portfolio attributes, such as fund age and the total number of funds in the same fund family. Khorana and Servaes (2012) find that product differentiation through starting new funds or innovation is an effective strategy to acquire market share. Khorana and Servaes (1999) find that families open new funds when the potential to generate additional fee income is substantial. Building on this, I hypothesize that product development through the base channel increases the market share and revenue of the family. The two sub-hypotheses within the second hypothesis are presented below. The first is the performance hypothesis and the second is the market share hypothesis:

**Hypothesis 2a: Fund family product quality development increases the performance in the family's funds.**

**Hypothesis 2b: Fund family product base development increases the market share of the family.**

In the next section, I explain the construction of the sample, present summary statistics, and explain the empirical methodology.

[18] As mentioned in Section 2.2, the quality index only contains funds with complete data within a year, i.e., funds that are started or liquidated during a year do not influence this index.

# 3 Data, summary statistics, and empirical methodology

This section presents the data and how the sample is constructed. Next, I present summary statistics for the sample, as well as summary statistics for input variables in the product development indices. The empirical methodology is explained in the last subsection.

## 3.1 Data

The main data source is Morningstar Direct. The starting point for the sample is constructed by downloading fund data on long-only open-ended equity mutual funds from 40 domiciles.[19] Index funds, enhanced index funds, ETFs, and funds-of-funds are excluded from the sample. After these exclusions, I am left with a sample consisting of 40,788 actively managed funds from 5,258 fund families. More details on the sample selection are presented in Appendix B.1. I download cross-sectional fund information, monthly gross returns, net returns, net assets, and fund size for all share classes from January 2006 to December 2019.[20] All of the time series variables are converted into USD in Morningstar Direct. I download country-level macro variables from the data library of the World Bank, as well as CPI data.[21] All dollar values are CPI-adjusted to the dollar-level as of December 2019.

The funds are categorized using the field *Morningstar Category*, and I make use of the MPT (Modern Portfolio Theory) index provided by Morningstar to compute benchmark-adjusted returns for the funds. This index is assigned by Morningstar, based on the portfolios of the funds, i.e., the actual investment universe, rather than the funds' self-reported benchmarks.[22] In addition to using MPT benchmark returns, I risk-adjust the returns using data from Ken French's data library and compute alpha estimates using CAPM, the Fama French 3-factor model, and the Fama French 3-factor model plus momentum.[23]

Next, the fund data is annualized. Annual returns are computed as the cumulative monthly returns within a year. In this process, funds with incomplete return data, i.e., less than 12 months of data within a year, are not given an annual return in these years to avoid bias in the return data from incomplete fund-year observations. Static variables are collected at the 31st of December each year. Finally, I use the fields *Domicile* and *Fund Company ID* to aggregate the fund data at the fund family-level to construct the family-year data sample. More details on the construction of the fund and fund family sample are provided in Appendix B.2. Fund-, family-, and industry-level control variables are explained in Appendix B.3.

To construct the product quality development index, I use the following fund data: gross returns, MPT benchmark returns, the field *Manager History* to detect manager changes, and *Morningstar Category* to define star funds. To construct the product base development index, I use the following fund data: the fields *Inception Date* to define fund starts, *Obsolete Date* to define fund liquidations,

[19] The domiciles include offshore domiciles, such as Luxembourg, Liechtenstein, and Ireland. Domiciles with less than 50 funds in the Morningstar database are excluded from the sample.

[20] Most of the funds are structured with multiple share classes. In Appendix B.2, I explain how the data is aggregated from share class-level to fund-level.

[21] Data available at: World Bank Data Library.

[22] More information about the MPT benchmark can be found at: MPT Benchmarks.

[23] The regional factors (Asia Pacific ex Japan, Developed, Emerging, Europe, Japan, North-America) are tied to the funds manually based on Morningstar Categories. The factors follow the computation of Fama and French (1993) and Fama and French (1996). Data available at: Ken French Data Library.

*Obsolete Type* and *Merged into Security ID* to define within-family mergers, and *Morningstar Category* to define new and liquidated fund categories.

Before starting the analysis, some restrictions are imposed on the sample. First, as explained in Section 2.2, by the definition of the product development indices, they cannot be computed for families with fewer than two funds. Thus, these families and their funds are excluded from the final sample. In addition, I am unable to compute the product development indices for families with missing data for any of the input variables. Finally, in order to perform a panel data study, I need a sufficient number of fund-year and family-year observations. Therefore, funds and fund families with less than three year-observations are excluded from the final sample. The next section presents summary statistics for the final sample of funds, fund families, and industries, as well as summary statistics for the input variables in the product development indices.

### 3.2 Summary statistics

This section presents summary statistics for the sample of funds, fund families, and industries, as well as the input variables in the product development indices. Table 1 reports general summary statistics, with fund characteristics in Panel I, fund family characteristics in Panel II, and industry characteristics in Panel III. All variables are of annual frequency. The summary statistics are from the sample after filtering in accordance with the description in the previous section and Appendix B.1.

The fund characteristics include the fund-level control variables and net flow, as well as performance measures, including the factor-adjusted alpha estimates. To ensure that extreme values do not drive the results in the fund sample, fund net flows and performance variables are winsorized at the bottom and top 1% level of the distribution across the whole sample. Even though the net flows are winsorized, there are still some outliers at the top of the distribution, which is reflected in the large maximum annual net flow and in the mean net flow of almost 14% of fund TNA at the end of the previous year. I also note that the average adjusted returns are negative for all of the factor-adjusted returns, and slightly positive for the MPT benchmark-adjusted returns.

For fund families, I present summary statistics for the control variables and fund family market share. The sample drops from around 200,000 fund-year observations to around 20,000 family-year observations. This shows that many of the funds in the sample are members of large fund families, with an average family size of around 11 funds. The average family market share is fairly high, at around 2%. The industry characteristics in Panel III include the industry competition variables explained in Section 2.1, as well as the other industry-level control variables. Overall, the table shows that there are large variations in the data, which is as expected for cross-country samples with large cross-sectional differences in the nature of the industries.

Next, I present summary statistics for the input variables in the product development indices in Table 2. The first two columns present the mean and standard deviation of the input variables, while the rest of the table contains the correlation matrix of the input variables. The correlations between most of the index components are low. The average correlations within the two subgroups (quality and base) are positive, which shows that these variables are to some extent related to each other. The correlation coefficient between the fund family average tracking error and fund family performance dispersion is high (0.54) compared to the other correlation coefficients, despite them measuring two

| | Panel I: Fund characteristics | | | | | |
|---|---|---|---|---|---|---|
| | Observations | Min | Median | Mean | Max | SD |
| Fund age (years) | 196,045 | 0.00 | 9.50 | 11.71 | 108.92 | 9.75 |
| TNA (million USD) | 189,999 | 0.00 | 82.15 | 570.23 | 236,685.20 | 3236.45 |
| Fund net flow (% of TNA) | 190,038 | -78.79 | -7.20 | 13.86 | 854.61 | 108.54 |
| Expense ratio (in %) | 178,442 | 0.20 | 1.53 | 1.64 | 210.07 | 1.01 |
| Return (in %) | 179,910 | -54.82 | 9.24 | 8.24 | 78.51 | 24.76 |
| MPT adjusted return (in %) | 173,461 | -22.20 | 0.05 | 0.35 | 26.94 | 7.85 |
| CAPM adjusted return (in %) | 158,992 | -25.64 | -0.52 | -0.30 | 29.14 | 8.79 |
| FF3 adjusted return (in %) | 158,992 | -24.27 | -0.50 | -0.42 | 25.95 | 7.93 |
| FF3MOM adjusted return (in %) | 158,992 | -23.30 | -0.44 | -0.29 | 25.69 | 7.79 |
| | **Panel II: Fund family characteristics** | | | | | |
| | Observations | Min | Median | Mean | Max | SD |
| Family age (years) | 19,120 | 1.00 | 17.67 | 20.68 | 95.42 | 15.19 |
| Family TNA (million USD) | 18,940 | 1 | 728 | 6,241 | 1,183,614 | 35,617 |
| Number of funds | 19,120 | 2.00 | 6.00 | 11.66 | 197.00 | 15.39 |
| Market share (in % of industry) | 18,940 | 0.00 | 0.35 | 2.17 | 87.20 | 5.33 |
| | **Panel III: Industry characteristics** | | | | | |
| | Observations | Min | Median | Mean | Max | SD |
| Industry TNA (million USD) | 560 | 401 | 27,344 | 225,101 | 6,854,653 | 870,251 |
| GDP per capita (thousand USD) | 529 | 807 | 40,542 | 40,648 | 178,846 | 30,407 |
| $Comp^{HHI}$ (%) | 557 | 20.74 | 87.67 | 84.32 | 97.46 | 12.43 |
| $Comp^{NHHI}$ (%) | 557 | 21.46 | 90.66 | 86.90 | 98.00 | 12.07 |

**Table 1. Sample summary statistics**
This table presents fund, fund family, and industry summary statistics of annual observations over the sample period of 2006 through 2019. Panel I presents summary statistics for the sample of funds, with fund-year observations, Panel II presents summary statistics for the sample of fund families, with family-year observations, while Panel III presents summary statistics for the sample of industries, with industry-year observations. Static variables are collected on the 31st December each year, and flow variables are annualized. Variables in USD are CPI-adjusted to the dollar-level as of December 2019.

different aspects of quality.[24] Some of the correlations across the subgroups are positive, but, in general, with low correlations. After aggregating the different components into the quality and base indices, the correlation between them is not statistically significant. This implies that the indices are essentially independent of each other, i.e. it appears to be no clear tradeoff between a quality development and base development strategy at the family-level. Thus, the choice to perform product quality (base) development appears to be independent of the choice to perform product base (quality) development.

### 3.3 Empirical methodology

This section outlines the main regressions in the empirical method. First, I present the fund family-level regression used to test the first hypothesis. To test the second hypothesis, I run both fund- and family-level regressions. Since the regressions are run on an international sample of funds and fund families, I add control variables for all levels equal to or below the level of the dependent variable. Control variables at the different levels are defined in Appendix B.3.

[24] As explained in Section 2.2, the tracking error is the average standard deviation of the benchmark-adjusted returns in the family, while the performance dispersion is the cross-sectional standard deviation of benchmark-adjusted fund returns within the family.

| | Mean | Sdev. | Tracking Error | Manager Change | Dispersion | Star Fund | Fund Start | Fund Liquid. | Fund Merg. | Category Start | Category Liquid. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Product Quality** | | | | | | | | | | | |
| Tracking Error (mean %) | 2.03 | 1.27 | 1.000 | | | | | | | | |
| Manager Change (% of funds) | 10.33 | 25.17 | 0.017 | 1.000 | | | | | | | |
| Dispersion (%) | 6.21 | 5.52 | 0.504 | 0.016 | 1.000 | | | | | | |
| Star Funds (% of funds) | 10.03 | 21.96 | 0.098 | -0.036 | 0.144 | 1.000 | | | | | |
| **Product Base** | | | | | | | | | | | |
| Fund Start (% of funds) | 0.17 | 3.16 | -0.010 | -0.009 | -0.012 | -0.005 | 1.000 | | | | |
| Category Liquid. (% of funds) | 0.06 | 1.81 | -0.002 | 0.007 | -0.003 | -0.011 | 0.001 | 1.000 | | | |
| Fund Mergers (% of funds) | 0.20 | 3.39 | 0.012 | 0.012 | 0.005 | -0.012 | 0.007 | 0.006 | 1.000 | | |
| Category Start (% of funds) | 7.24 | 28.00 | 0.063 | 0.028 | 0.027 | -0.067 | 0.022 | 0.004 | -0.007 | 1.000 | |
| Category Liquid. (% of funds) | 0.63 | 6.66 | 0.114 | -0.009 | 0.095 | -0.009 | 0.001 | 0.026 | 0.007 | 0.010 | 1.000 |

**Table 2. Summary statistics for product development input variables**
This table presents the family-level summary statistics for the input variables in the product development indices, as well as the correlations between the input variables.

### Hypothesis 1

To test the first hypothesis concerning the relationship between competition and fund family product development, I run family-year level panel regressions, controlling for industry and family variables. The main test variables are the industry competition variables explained in Section 2.1, while the dependent variables are the product development indices derived in Section 2.2. The baseline regression is given by:

$$Index_{f,c,t} = \alpha_{f,c} + \alpha_t + \gamma Comp_{f,c,t-1} + \beta_1 X_{f,c,t-1} + \beta_2 X_{c,t-1} + \epsilon_{f,c,t} \tag{9}$$

where $\alpha_{f,c}$ are fund family fixed effects and $\alpha_t$ are year fixed effects. $Comp_{f,c,t-1}$ is lagged industry competition, $X_{f,c,t-1}$ is a vector containing lagged family-level covariates, and $X_{c,t-1}$ is a vector containing lagged industry-level covariates. The control variables are lagged to minimize potential issues with endogeneity, i.e., I test the how the industry competition at the end of the previous year affects fund family product development in the following year.[25]

There is a growing discussion in the mutual fund literature about the use of fixed effects and cluster specifications in panel regressions (see, e.g., Pástor, Stambaugh and Taylor (2017)). In my case, the variation of interest is the within-family variation in product development indices ($Index$) in response to variation in the competition variable ($Comp$). Thus, I include fund family fixed effects to control for this variation in the regressions. In addition, I control for potential year effects by including year fixed effects. I also impose two different cluster specifications. I account for within-family autocorrelation by clustering the standard errors along families. Next, the standard errors are independently clustered along both family and time, to also account for potential correlations across families within years.

### Hypothesis 2

I run two types of regressions to test the second hypothesis. The first are fund-level regressions with fund alpha estimates and net flow as dependent variables. The second are family-level regressions with family market share as dependent variable. In both types of regressions, the main test variables are the product development indices. The baseline fund-level regression is on the following form:

$$y_{i,f,c,t} = \alpha_{i,f,c,t} + \alpha_t + \gamma Index_{f,c,t-1} + \beta_1 X_{i,f,c,t-1} + \beta_2 X_{f,c,t-1} + \beta_2 X_{c,t-1} + \epsilon_{i,f,c,t} \tag{10}$$

[25] I elaborate more on this when interpreting the results in Section 4.

where $\alpha_{i,f,c,t}$ are fund fixed effects and $\alpha_t$ are year fixed effects. The dependent variables are fund performance, adjusted both without and with factor models, and fund net flow. $Index_{f,c,t-1}$ is the lagged product development index, where I run separate regressions testing each index. $X_{i,f,c,t-1}$ is a vector containing lagged fund-level covariates, $X_{f,c,t-1}$ is a vector containing lagged family-level covariates, and $X_{c,t-1}$ is a vector containing lagged industry-level covariates.

Next, the regressions testing how fund family product development affects the family market share are on the form:

$$MS_{f,c,t} = \alpha_{f,c,t} + \alpha_t + \gamma Index_{f,c,t-1} + \beta_2 X_{f,c,t-1} + \beta_2 X_{c,t-1} + \epsilon_{f,c,t} \tag{11}$$

where the dependent variable, $MS_{f,c,t}$, is the market share of the family. I run these regressions both with and without control variables, since the market share variable does not contain a lot of variation from year to year.

# 4 Tests of hypothesis 1: competition and product development

This section tests the first hypothesis. I hypothesize that fund families increase product development when the competitive pressure increases. The empirical methodology comprises the regression explained in Equation (9), with the industry competition variables from Section 2.1 as test variables, and the product development indices from Section 2.2 as dependent variables. A typical concern when dealing with industry competition variables is reverse causality, i.e., that industry competition is greater in industries where fund family product development is higher, and vice versa. If this is the case, the results from the standard panel regressions might be biased. Therefore, I also extend the test by proposing an instrumental variable approach to deal with this.

## 4.1 Main results

The regression in Equation (9) estimate the within-family variation in product development with the variation in industry competition. Relating the hypothesis test to Equation (9), the industry competition coefficients, $\gamma$, are interpreted as follows:

- Fund family product development does not change with the level of industry competition: $\gamma = 0$.
- Fund family product development increases with the level of industry competition: $\gamma > 0$.
- Fund family product development decreases with the level of industry competition: $\gamma < 0$.

I hypothesize that $\gamma > 0$ for both of the product development channels, i.e., that fund families carry out product development to differentiate themselves from competitors when competition increases. Table 3 reports the results, testing industry competition in Panel I, and normalized industry competition in Panel II. In both panels, the dependent variables in column (1)-(2) are the product quality development index, in column (3)-(4) the product base development index, in column (5)-(6) the total product development index, and in column (7)-(8) the product quality development index net of the product base development index. The regressions for each product development index are first run with standard errors clustered along family and then with standard errors independently clustered along both family and time.

Starting with the results testing industry competition, I find that the coefficient of $Comp_{t-1}^{HHI}$ is positive and significant at the 5% level in the first regression on the product quality development index and negative and significant at the 10% level in the first regression on the product base development index. The independent two-way clustered standard errors are higher and the statistical significance declines, but the results are still statistically significant at the 10% level in both tests. The magnitude of the competition coefficient is higher in the regressions on the product base development index, which also gives a negative, but statistically insignificant competition coefficient in the regressions with total product development as dependent variable. The opposite nature of the coefficients in the regressions on product quality and product base development lead to a positive and highly statistically significant coefficient in the regressions with the net quality development index as dependent variable. The results in Panel II are similar, but the statistical significance in the tests of product quality development increases. As explained in Section 2.1, the normalized industry competition corrects for the number of firms in the industry. This shows that the number of firms does not influence the results in Panel I, and that $Comp^{HHI}$ measures industry competition, and not the number of firms in the industry.

These results show that greater competition is associated with an increase in product quality development in fund families. I also find some evidence that increased industry competition is associated with a reduction in product base development. For total product development, I find no significant effect of competition, but I find that the net quality development increases with greater competition. Since the decile ranks are computed and normalized in the entire cross-section each year, they measure changes in the input variables relative to the other families in the cross-section. In terms of economic significance, a one within-family average standard deviation increase in $Comp^{HHI}$ (0.011) corresponds to an increase of 0.0072 ($0.011 \cdot 0.652$) in the family product quality development index. Compared to the within-family average product development index of $-0.026$, this corresponds to an increase of around 27% in the quality development index. The results testing $Comp^{NHHI}$ have similar economic significance. Correspondingly, a one standard deviation increase in $Comp^{HHI}$ is associated with a reduction of 0.0078 ($0.011 \cdot -0.710$), or 18% compared to the sample mean, in the product base development index.

While the main focus of these tests is the time series relationship between industry competition and fund family product development, I also run cross-sectional first-difference regressions in Appendix C.1. Table C1 reports country-level regressions, while Table C2 reports family-level regressions. I find that the relationship between competition and product quality development also holds in the cross-section, in both regression specifications. Moreover, I find negative and statistically significant coefficients in the tests of product base development, which correspond to the opposite nature of the relationship to industry competition in the two channels from the main tests. Next, as a robustness test to examine whether the results hold globally, I run the main tests on a sample excluding the US. The US is well-known to have the most developed mutual fund industry in the world.[26] The results are presented in Table C3 in Appendix C.2. The importance of the US in the total sample is illustrated by the number of observations in these regressions, dropping by around 1/4 relative to the full sample. The statistical significance of the results holds in the sample excluding the US, with a minor decline in magnitude. Based on these results, I argue that the findings in this section are not driven by the mechanics or competitiveness of the US mutual fund industry.

The results from the family fixed-effects regressions show that the families respond to greater competition with different strategies in the two channels. While they scale up product quality development, they also scale down product base development. While the correlation between the indices are not statistically significant, as noted in Section 3.2, these results suggest that, in order to scale up product quality development, the families also reduce product base development. One reason for this might be that it is costly to increase quality development, and that the families have to scale down product base development to do so. As mentioned in the introduction to this section, reverse causality is a potential concern in these tests, and the results could consequently be biased. To examine this further, I propose an instrumental variable (IV) approach.

[26] In the fund family sample, the assets under management in the US industry account for around 60% of the total assets under management. See, for example, Investment Company Institute (2019) for more information about the size of the US mutual fund industry.

| | Panel I: Industry competition | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | Product quality development | | Product base development | | Product development | | Net product quality development | |
| $Comp_{t-1}^{HHI}$ | 0.652** | 0.652* | −0.710* | −0.710* | −0.058 | −0.058 | 1.363*** | 1.363** |
| | (0.268) | (0.331) | (0.393) | (0.380) | (0.443) | (0.359) | (0.506) | (0.615) |
| Controls | Family and Industry | | | | | | | |
| FE | Family + Year | | | | | | | |
| Cluster | Family | Family + Year | Family | Family + Year | Family | Family + Year | Family | Family + Year |
| Obs. | 17,110 | 17,110 | 17,110 | 17,110 | 17,110 | 17,110 | 17,110 | 17,110 |
| $R^2$ | 0.377 | 0.377 | 0.266 | 0.266 | 0.308 | 0.308 | 0.335 | 0.335 |
| Adj. $R^2$ | 0.298 | 0.298 | 0.173 | 0.173 | 0.221 | 0.221 | 0.251 | 0.251 |
| | **Panel II: Normalized industry competition** | | | | | | | |
| Dep. var. | Product quality development | | Product base development | | Product development | | Net product quality development | |
| $Comp_{t-1}^{NHHI}$ | 0.674*** | 0.674** | −0.629* | −0.629 | 0.045 | 0.045 | 1.304*** | 1.304** |
| | (0.256) | (0.309) | (0.381) | (0.372) | (0.428) | (0.334) | (0.488) | (0.596) |
| Controls | Family and Industry | | | | | | | |
| FE | Family + Year | | | | | | | |
| Cluster | Family | Family + Year | Family | Family + Year | Family | Family + Year | Family | Family + Year |
| Obs. | 17,110 | 17,110 | 17,110 | 17,110 | 17,110 | 17,110 | 17,110 | 17,110 |
| $R^2$ | 0.377 | 0.377 | 0.266 | 0.266 | 0.308 | 0.308 | 0.335 | 0.335 |
| Adj. $R^2$ | 0.298 | 0.298 | 0.173 | 0.173 | 0.221 | 0.221 | 0.251 | 0.251 |

**Table 3. Family product development and competition regressions**
Regressions testing the effect of industry competition on fund family product development indices, on the form presented in Equation (9). The product development indices are presented in Section 2. Panel I tests industry competition based on *HHI*, and Panel II tests normalized industry competition based on *NHHI*, both explained in Section 2.1. All regressions include lagged controls at the family- and industry-level. Control variables are described in Appendix B.3. Cluster robust standard errors are reported in the parentheses, with the level of clusters reported in the table. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

## 4.2 Instrumental variable results

Reverse causality is a problem in the sense that industry competition might be greater in industries where fund family product development is higher, and vice versa. In the main regressions, I lag the control variables to the end of the previous year to distinguish between the variation in the test variables and the effect of this variation on the dependent variable. However, when the variation in the industry competition variable is low, fund families might form expectations based on industry competition in the previous year. Therefore, to formally deal with the potential endogeneity issues, I instrument the industry competition and run instrumental variable regressions. Following Rodríguez-Castelán et al. (2019), I treat *HHI* and *NHHI* as endogenous and instrument them using Bartik-style share-shift instruments (see Bartik (1991) and Bartik (2002)).

Bartik instruments are designed to reduce reverse causality and correlated omitted variable concerns in differential exposure designs (e.g., panel regressions with unit and time effects). To isolate plausibly exogenous variation in industry competition, and to use the fact that trends in industry competition impact industries differently, the instruments are constructed using two different growth rates. First, I use the growth in *HHI* and *NHHI* for the US active mutual fund industry, and, second, I use the global value-weighted growth in *HHI* and *NHHI*. More details on the construction of the instruments are provided in Appendix C.3. To go from the instrumented industry concentration variables to the instrumented industry competition variables, the instrumented industry concentration is subtracted from 1, as explained in Section 2.1.

The change in industry concentration can then be divided into an "exogenous component" recreated using the instrument and deviations from this predicted change, and the instruments allow for using only the exogenous source of variation. By design, the instrument isolates the component of change in the local industry that is driven by US or global trends, such as change in demand for actively managed mutual funds or other industry trends.

Table 4 reports the results for instrumented industry competition (Panel I) and instrumented normalized industry competition (Panel II). Both panels report the results using US growth as instrument in the five first columns and global growth as instrument in the last five columns. The regressions using the US growth rate as instrument does not include observations from the US. All first stage regressions have highly statistically significant coefficients of the instrument on the observed competition variable, i.e., the relevance condition holds. In the regression tests, I find statistically significant coefficients of the instrumented competition on the product quality development index. Comparing the coefficients from the tests of the product quality development index to the corresponding coefficients in Table 3, they are generally around four times larger, while the instrumented industry competition has 50-80% of the standard deviation compared to the observed industry competition. Thus, these tests suggest that the true effect of competition on product quality development is at least twice as large as the estimates from the previous section. Furthermore, in the IV-results the industry competition coefficients are positive, but not statistically significant, for regressions with the product base development index as a dependent variable. These tests suggest that the main results are biased.

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Panel I: Instrumented industry competition** | | | | | | | | | | |
| Instrument | | | US growth | | | | | Global growth | | |
| Dep. var. | First Stage | Quality | Base | Development | Net quality | First Stage | Quality | Base | Development | Net quality |
| $Comp^{Bartik}_{HHI}$ | 0.494*** | | | | | 0.834*** | | | | |
| | (0.106) | | | | | (0.067) | | | | |
| $Comp^{IV}_{HHI}$ | | 2.590** | 1.330 | 3.920 | 1.261 | | 2.305*** | 0.442 | 2.747 | 1.863* |
| | | (1.259) | (1.715) | (2.557) | (1.587) | | (0.777) | (1.167) | (1.707) | (1.008) |
| Controls | Family + Industry | | | | | | | | | |
| FE | Family + Year | | | | | | | | | |
| Cluster | Family | | | | | | | | | |
| Obs. | 13,416 | 13,416 | 13,416 | 13,416 | 13,416 | 17,110 | 17,110 | 17,110 | 17,110 | 17,110 |
| $R^2$ | 0.937 | 0.372 | 0.272 | 0.314 | 0.328 | 0.943 | 0.376 | 0.266 | 0.307 | 0.335 |
| Adj. $R^2$ | 0.929 | 0.290 | 0.177 | 0.224 | 0.240 | 0.936 | 0.296 | 0.172 | 0.219 | 0.251 |
| **Panel II: Instrumented normalized industry competition** | | | | | | | | | | |
| Instrument | | | US growth | | | | | Global growth | | |
| Dep. var. | First Stage | Quality | Base | Development | Net quality | First Stage | Quality | Base | Development | Net quality |
| $Comp^{Bartik}_{NHHI}$ | 0.526*** | | | | | 0.857*** | | | | |
| | (0.104) | | | | | (0.065) | | | | |
| $Comp^{IV}_{NHHI}$ | | 2.297** | 1.359 | 3.656 | 0.939 | | 2.094*** | 0.564 | 2.658* | 1.530 |
| | | (1.116) | (1.614) | (2.336) | (1.498) | | (0.697) | (1.130) | (1.604) | (0.975) |
| Controls | Family + Industry | | | | | | | | | |
| FE | Family + Year | | | | | | | | | |
| Cluster | Family | | | | | | | | | |
| Obs. | 13,416 | 13,416 | 13,416 | 13,416 | 13,416 | 17,110 | 17,110 | 17,110 | 17,110 | 17,110 |
| $R^2$ | 0.922 | 0.372 | 0.272 | 0.314 | 0.328 | 0.928 | 0.376 | 0.265 | 0.307 | 0.335 |
| Adj. $R^2$ | 0.912 | 0.291 | 0.177 | 0.225 | 0.240 | 0.919 | 0.297 | 0.172 | 0.219 | 0.251 |

**Table 4. Family product development and competition IV-regressions**
Instrumental variable regressions testing the effect of competition on fund family product development indices, on the form presented in Equation (9). The industry competition variables are instrumented using US growth and global growth in industry competition, as explained in Appendix C.3. The product development indices are presented in Section 2. Panel I tests instrumented industry competition based on *HHI*, and Panel II tests instrumented normalized industry competition based on *NHHI*, both explained in Section 2.1. All regressions include lagged controls at the family- and industry-level. Control variables are described in Appendix B.3. Cluster robust standard errors are reported in the parentheses, with the level of clusters reported in the table. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

In both the main regressions and the IV-regressions, I find that greater industry competition is associated with an increase in product quality development. In addition, I find some evidence in the main tests and the cross-sectional first-difference regressions of product base development declining with increased competition, but this does not hold when I introduce instrumental variables. Also, in the IV-results, the estimated effects of competition on product quality development increase and the sign of the effects on product base development changes from negative to positive. The IV-results highlight the importance of accounting for endogeneity in tests of how industry competition affects firm-level variables. In sum, these results show that the effect of competition on product quality development is larger than I estimated in the panel regressions, and that increasing industry competition does not appear to have a statistically significant effect on product base development. The difference in the product quality and product base development responses to competition is an interesting finding, as it shows that fund families prefer to increase the quality of their funds rather than expanding their fund base when competition increases. In the test of the second hypothesis, I examine how these two different channels impact the performance and market share outcomes of fund families.

# 5 Tests of hypothesis 2: product development outcomes

This section tests the second hypothesis concerning how fund family product development strategies affect fund performance and fund family market share outcomes. The dependent variables in the regressions are defined in the text. I hypothesize that product quality development increases performance in family-affiliated funds and that product base development increases the revenue of the family, i.e., that the two channels of product development make the families more competitive in two different ways.

## 5.1 H2a: family product development and fund performance

This section tests the first part of the hypothesis, i.e., how product development strategies affect performance or value creation in family-affiliated funds. Multiple studies find that fund families can impact and affect the performance of their funds (see, e.g., Guedj and Papastaikoudi (2003) and Gaspar et al. (2006)). Fang et al. (2014) and Berk et al. (2017) find that fund families strategically allocate their fund managers to maximize the total value creation in the family. To examine how fund family product development affects fund performance, I first define how to measure fund performance. Benchmark-adjusted returns, often called gross alpha, is a standard performance measure in the mutual fund literature. To compute benchmark-adjusted returns, I make use of the MPT benchmark provided by Morningstar. However, the downside of using benchmark-adjusted returns is the implicit assumption that the beta-exposures to the benchmark are equal to one for all funds. In other words, this measure does not take the risk of the fund into account. Therefore, in addition to using the MPT benchmarks to adjust the returns, I also estimate risk-adjusted returns using factor returns from Ken French's website. The factor return data is described in Section 3.1.

The risk-adjusted returns are estimated in two steps. First, I run factor regressions for each fund with monthly data over the full data period of the fund, and extract the factor coefficients.[27] The alpha estimates are estimated using three different factor models: the CAPM model, the Fama French 3-factor model (referred to as FF3) adding the size factor (*SMB*) and the value factor (*HML*), and the Fama French 3-factor plus the momentum factor (*WML*) (referred to as FF3MOM).[28] Illustrated by the latter model, which includes all potential factors, I run factor regressions on the following form for all funds:

$$R_{i,t} - R_{f,t} = \alpha + \beta(R_{m,t} - R_{f,t}) + \beta_{SMB}SMB + \beta_{HML}HML + \beta_{WML}WML + \epsilon_{i,t} \qquad (12)$$

where $R_{i,t}$ is the gross return of the fund and $R_{f,t}$ is the risk-free rate. In the next step, the factor coefficient estimates are used together with the returns of each factor to risk-adjust the monthly fund returns. Finally, the monthly data is annualized in the same way as the fund returns described in Section 3.1.

Table 5 presents cross-sectional differences in the performance of funds from families with high and low development in the two product development channels.[29] The portfolios are formed at the end

[27] For statistical robustness purposes, I exclude funds with less than 24 months of return observations.

[28] The size factor *SMB* is the return of small stocks minus the return of large stocks, while the value factor *HML* is the return of stocks with a high book-to-market ratio minus the return of stocks with a low book-to-market ratio. The momentum factor *HML* is the return of stocks with high returns minus stocks with low returns in the past 12 months.

[29] High product development is defined as the top 20th (Q5) percentile and low product development as the bottom 20th (Q1) percentile within each domicile in each year.

of year $t-1$, and the performance figures are annual alpha estimates for the following year $t$. The differences between the performance in the top and bottom 20th percentile based on product quality development are statistically significant, with the highest alpha in the top 20th percentile. These results hold for all of the alpha estimates, which shows that the difference cannot be attributed to differences in risk exposure. I also note that the difference decreases for risk-adjusted alpha, and that all of the risk-adjusted alpha estimates are negative.[30] For product base development portfolios, the differences in risk-adjusted performance are for the most part statistically insignificant, and the alpha in these portfolios are lower than for the top 20th percentile based on the product quality index.

| | **Product quality development index** | | | | **Product base development index** | | | |
|---|---|---|---|---|---|---|---|---|
| Percentile | Top 20th | Bottom 20th | Diff. | P-val. diff. | Top 20th | Bottom 20th | Diff. | P-val. diff. |
| MPT-adjusted | 0.60 | -0.05 | 0.65 | 0.00 | 0.34 | 0.18 | 0.17 | 0.01 |
| CAPM-adjusted | -0.15 | -0.82 | 0.68 | 0.00 | -0.48 | -0.43 | -0.05 | 0.52 |
| FF3-adjusted | -0.36 | -0.90 | 0.54 | 0.00 | -0.59 | -0.58 | -0.01 | 0.91 |
| FF3MOM-adjusted | -0.23 | -0.66 | 0.43 | 0.00 | -0.48 | -0.41 | -0.07 | 0.29 |

**Table 5. Performance in product development portfolios**
Performance of the top and bottom 20th percentile portfolios based on the product development indices. The alpha estimates are MPT benchmark-adjusted returns, CAPM-adjusted returns, FF3-adjusted returns, and FF3MOM-adjusted returns. The differences are the returns in the top 20th percentile minus the bottom 20th percentile, and the p-values are from T-tests testing mean equal to zero.

Next, I formally test the relationship between product development and performance by running the fund-level regressions from Equation (10). The regressions include fund-, family-, and industry-level control variables, as well as the lagged alpha, estimated in the same way as the dependent variable in each regression. As explained in the hypotheses derivation in Section 2, the objective is to test whether product development makes the fund families more competitive in terms of performance. Relating the hypothesis test to Equation (10), I interpret the regression coefficients of the product development index variables, $\gamma$, as follows:

- Product development does not affect fund performance: $\gamma = 0$.
- Product development increases fund performance: $\gamma > 0$.
- Product development reduces fund performance: $\gamma < 0$.

The initial hypothesis is that $\gamma > 0$ for the product quality development index, i.e., that product quality development increases the performance of family-affiliated funds. I also run regressions testing the product base development index, as well the combined indices. Table 6 reports the regression results for MPT benchmark-adjusted returns (Panel I), CAPM-adjusted returns (Panel II), FF3-adjusted returns (Panel III), and FF3MOM-adjusted returns (Panel IV). The regressions testing each product development index are first run with standard errors clustered along fund and next with standard errors independently clustered along fund and time.

I find positive and statistically significant coefficients for the product quality development index in the regressions with fund-level clustered standard errors. For all alpha estimates, the statistical

[30]This is in line with findings from previous studies showing that the active mutual fund industry struggles to add value as a whole.

significance drops when the standard errors are independently clustered along both fund and time, but some of the product quality development coefficients still hold at the 5 and 10% level.[31] In the regressions testing product base development on fund performance, I find no evidence for this channel affecting the performance of the funds. For the total product development index and net quality development index, I find similar results as for the product quality development index, which shows that these results are driven by the effect of the product quality development index. Table D1 in Appendix D.1 reports corresponding regressions with value added estimates as dependent variables. Value added is is an estimate of how much value the fund extracts from the financial markets, measured in USD (Berk and Van Binsbergen (2015)).[32] I find that product quality development increases value added in the funds, but that the statistical significance is lower compared to the tests of alpha estimates in Table 6. Furthermore, I find the same difference in how the two channels affect the fund's value creation.

The average within-fund standard deviation of the product quality development index is 0.64. Thus, a one standard deviation increase in the fund family product quality development index corresponds to an increase of 0.13% ($0.64 \cdot 0.199\%$) in the annual MPT-adjusted returns, and of 0.07% ($0.64 \cdot 0.114\%$) in the annual FF3-adjusted returns, which is the lowest estimated effect on alpha. These numbers do not appear to be large, but compared to the corresponding within-fund average annual alpha estimates of 0.16% and $-0.65\%$, respectively, I argue that product development appears to have a relatively large effect on the MPT-adjusted returns. Correspondingly, the contribution of product quality development to value added estimates ranges from 0.47 ($0.64 \cdot 0.741$) million USD for the MPT-adjusted value added to 0.11 ($0.64 \cdot 0.171$) million USD for the FF3MOM-adjusted value added. The within-family sample averages are 1.25 and $-0.85$. Thus, the contribution of product quality development also accounts for a fairly large share of the MPT benchmark-adjusted value added.

The results in this sections support the hypothesis that product quality development increases performance. The effects on alpha production are fairly large compared to the sample mean. A one standard deviation increase in the product quality development index accounts for around 80% (0.13/0.16) of the unconditional sample average MPT benchmark-adjusted returns. The results for value added estimates are consistent with the tests of alpha estimates. I also find differences in the effects of product quality and product base development on the performance of the family-affiliated funds. In sum, these results show that fund family efforts made to increase the quality of the funds have a positive effect on value creation in the funds. They also show that fund family product development strategies can affect the performance of the family's funds.

### 5.2 H2b: family product development and market share

Next, I test the second sub-hypothesis within hypothesis 2. As explained in the hypothesis derivation in Section 2.3, fund families can adopt different strategies for differentiating themselves from rival families, based on the product development indices. The first is to make efforts to increase the overall performance of their funds, while the second is to make efforts to increase market share and revenue. In this section, I examine how the product development channels affect the market share of the fund

[31] This double cluster is more conservative, as it also corrects for the within-year correlation across funds (see, for example, the discussion in Cremers, Fulkerson and Riley (2019)).

[32] See Appendix D for details on how value added is computed.

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Panel I: MPT benchmark-adjusted returns** | | | | | | | | |
| Index | Quality | | Base | | Total | | Net quality | |
| $Index_{t-1}$ | 0.199*** | 0.199** | 0.016 | 0.016 | 0.099*** | 0.099** | 0.083*** | 0.083** |
| | (0.031) | (0.068) | (0.028) | (0.030) | (0.020) | (0.033) | (0.021) | (0.037) |
| Controls | Fund + Family + Industry | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 147,659 | 147,659 | 147,659 | 147,659 | 147,659 | 147,659 | 147,659 | 147,659 |
| $R^2$ | 0.214 | 0.214 | 0.214 | 0.214 | 0.214 | 0.214 | 0.214 | 0.214 |
| Adj. $R^2$ | 0.101 | 0.101 | 0.100 | 0.100 | 0.100 | 0.100 | 0.100 | 0.100 |
| **Panel II: CAPM-adjusted returns** | | | | | | | | |
| Index | Quality | | Base | | Total | | Net quality | |
| $Index_{t-1}$ | 0.136*** | 0.136 | −0.024 | −0.024 | 0.048* | 0.048 | 0.075*** | 0.075* |
| | (0.037) | (0.094) | (0.033) | (0.076) | (0.025) | (0.075) | (0.025) | (0.039) |
| Controls | Fund + Family + Industry | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 |
| $R^2$ | 0.222 | 0.222 | 0.222 | 0.222 | 0.222 | 0.222 | 0.222 | 0.222 |
| Adj. $R^2$ | 0.109 | 0.109 | 0.109 | 0.109 | 0.109 | 0.109 | 0.109 | 0.109 |
| **Panel III: FF3-adjusted returns** | | | | | | | | |
| Index | Quality | | Base | | Total | | Net quality | |
| $Index_{t-1}$ | 0.114*** | 0.114* | 0.036 | 0.036 | 0.071*** | 0.071 | 0.032 | 0.032 |
| | (0.033) | (0.057) | (0.030) | (0.062) | (0.022) | (0.045) | (0.022) | (0.040) |
| Controls | Fund + Family + Industry | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 |
| $R^2$ | 0.224 | 0.224 | 0.224 | 0.224 | 0.224 | 0.224 | 0.224 | 0.224 |
| Adj. $R^2$ | 0.112 | 0.112 | 0.112 | 0.112 | 0.112 | 0.112 | 0.112 | 0.112 |
| **Panel IV: FF3MOM-adjusted returns** | | | | | | | | |
| Index | Quality | | Base | | Total | | Net quality | |
| $Index_{t-1}$ | 0.134*** | 0.134* | 0.030 | 0.030 | 0.077*** | 0.077 | 0.044** | 0.044 |
| | (0.033) | (0.074) | (0.029) | (0.062) | (0.022) | (0.049) | (0.022) | (0.046) |
| Controls | Fund + Family + Industry | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 |
| $R^2$ | 0.220 | 0.220 | 0.220 | 0.220 | 0.220 | 0.220 | 0.220 | 0.220 |
| Adj. $R^2$ | 0.108 | 0.108 | 0.107 | 0.107 | 0.108 | 0.108 | 0.107 | 0.107 |

**Table 6. Family product development and fund performance regressions**

Regressions testing the effect of fund family product development indices on fund performance, on the form presented in Equation (10). The product development indices are presented in Section 2. The dependent variable in the regressions in Panel I is MPT-adjusted return, dependent variable in Panel II is CAPM-adjusted return, dependent variable in Panel III is FF3-adjusted return, and dependent variable in Panel IV is FF3MOM-adjusted return. All regressions include lagged controls at the fund-, family-, and industry-level, as well as lagged adjusted returns, adjusted in the same way as the dependent variable. Control variables are described in Appendix B.3. Cluster robust standard errors are reported in the parentheses, with the level of cluster presented in the table. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

families. Relating the hypothesis test to the regressions in Equation (10) and (11), the coefficients of the product development indices, $\gamma$, are interpreted as follows:

- Product development does not affect the market share of the fund family: $\gamma = 0$.

- Product development increases the market share of the fund family: $\gamma > 0$.
- Product development reduces the market share of the fund family: $\gamma < 0$.

I hypothesize that $\gamma > 0$ for the product base development index, i.e., that product base development increases the market share of the family. As in the previous section, I also run regressions testing the other product development indices. The hypothesis test is split into two different regressions. The first comprises fund-level regressions with net flow as the dependent variable, while the second comprises family-level regressions with family market share as the dependent variable. While the tests of market share test the hypothesis directly, the net flow regressions shows whether there are any spillover effects from product development at the family-level.

First, I define the net flow variables used in the fund-level regressions. The monthly net flow of a fund is defined as the increase in TNA in excess of the value increase in the underlying portfolio. For fund $i$ in month $t$, the net flow is computed as follows:

$$NF_{i,t} = TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t}). \tag{13}$$

The annual net flow is computed as the sum of monthly net flows within the year, divided by the assets under management in the fund at the end of the previous year. The regression results from the fund-level regressions in Equation (10) with net flow as dependent variable are presented in Table 7. The results show that product development through the product base channel increases the net flows to funds in the family. Development through the quality channel also has a positive, but statistically insignificant coefficient. Total product development also increases net flows, and this effect is driven by the effect of product base development. The results mean that a one standard deviation increase in the product base development index increases the fund net flows by 0.73% ($0.76 \cdot 0.959\%$). The unconditional average within-fund net flow is equal to 12.8%. Thus, the estimated effect is relatively low compared to the sample mean. As mentioned in Section 3.2, net flows are winsorized at the bottom and top 1% level of the distribution of the entire sample. However, with this level of winsorizing, there are still some large outliers at the top of the distribution, which is also reflected by the high average net flow in Table 1. As a robustness test to examine whether the extreme values drive the results, I run corresponding regressions in Table D2 in Appendix D.2 with net flows winsorized at the bottom and top 3% and 5% level of the distribution. In these samples, the average within-fund net flows are equal to 6.25% and 2.54%, respectively. With stricter winsorizing of net flows, I find that increasing product quality development index also increases the net flows to family-affiliated funds. This might be because product quality development increases performance, and that investors tend to chase past performance. However, the results for product base development are stronger than for product quality development in terms of both economic and statistical significance, which suggests that this channel is a better source of increasing the net flows to the family's funds.

Second, as main test of the market share hypothesis, I test how product development strategies affect the market share of the family directly. The regressions are on the form explained in Equation (11), with fund family market share as the dependent variable. The objective is to test whether product development makes the fund families more competitive in terms of market power. The market share

| | Net flow (%) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Index | Quality | | Base | | Total | | Net quality | |
| $Index_{t-1}$ | 0.183 | 0.183 | 0.959*** | 0.959** | 0.601*** | 0.601** | −0.438* | −0.438 |
| | (0.325) | (0.396) | (0.311) | (0.426) | (0.225) | (0.264) | (0.225) | (0.318) |
| Controls | Fund + Family + Industry | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 |
| $R^2$ | 0.326 | 0.326 | 0.326 | 0.326 | 0.326 | 0.326 | 0.326 | 0.326 |
| Adj. $R^2$ | 0.229 | 0.229 | 0.229 | 0.229 | 0.229 | 0.229 | 0.229 | 0.229 |

**Table 7. Family product development and fund net flow regressions**
Regressions testing the effect of fund family product development indices on fund net flow, on the form presented in Equation (10). The product development indices are presented in Section 2. The dependent variable is annual net flow in percentage of TNA at the end of the previous year. All regressions include lagged controls at the fund-, family-, and industry-level. Control variables are described in Appendix B.3. Cluster robust standard errors are reported in the parentheses, with the level of cluster presented in the table. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

of family $f$ in country $c$ at time $t$ is the family's TNA divided by the total industry TNA

$$MS_{f,c,t} = \frac{TNA_{f,c,t}}{TNA_{c,t}}. \tag{14}$$

Since the variation in the dependent variable is expected to be low, i.e., the dependent variable is highly persistent, I run the regressions both without and with control variables, where both specifications include family and time fixed effects. The regression results are presented in Table 8. I find positive and significant coefficients for the product base development index. When adding control variables, the statistical significance decreases from being statistically significant at the 1% level to the 10% level. As in the net flow regressions, I find positive coefficients of the product quality development index, but they are not statistically significant. The coefficient for total product development is positive and significant, driven by the product base development index. Furthermore, even when excluding control variables, the $R^2$ is high, which confirms that the variation in the dependent variable is low and can to a large extent be explained by family and time fixed effects, together with the product development indices.

The tests of the market share hypothesis show that product development through the base channel increases net flows to the funds in the family. The estimated effect is low compared to the average flow in the sample, but it shows how product base development can be an effective strategy for increasing the assets under management in a family. The increase in flows is also leads to an increase in the market share of the family. Moreover, the results show weak evidence for product quality development being a means of increasing the market share.

To test whether the estimated effects of the product development indices are in fact the result of coordinated fund family product development, I run placebo tests on samples of pseudo-families. The pseudo-families are constructed by matching funds based on domicile, fund style, and fund size. Thus, each fund is matched to a fund from another family, where the fund is in the same domicile, have the same style, and is from the same fund size decile. The matching is one-to-one, i.e., that each fund is only matched to one other fund. In the case of multiple valid matches, the fund with most similar TNA is used as match. I run two different matching procedures, where the difference is the definition

| | Fund family market share | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Index | Quality | | Base | | Total | | Net quality | |
| $Index_{t-1}$ | 0.012 | 0.014 | 0.052*** | 0.031* | 0.034*** | 0.023** | −0.024** | −0.011 |
| | (0.012) | (0.011) | (0.015) | (0.017) | (0.009) | (0.010) | (0.011) | (0.011) |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| FE | Fund + Year | | | | | | | |
| Cluster | Family | | | | | | | |
| Obs. | 16,740 | 16,226 | 16,740 | 16,226 | 16,740 | 16,226 | 16,740 | 16,226 |
| $R^2$ | 0.935 | 0.947 | 0.935 | 0.947 | 0.935 | 0.947 | 0.935 | 0.947 |
| Adj. $R^2$ | 0.928 | 0.941 | 0.928 | 0.941 | 0.928 | 0.941 | 0.928 | 0.941 |

**Table 8. Family product development and family market share regressions**
Regressions testing the effect of fund family product development indices on fund family market share, on the form presented in Equation (11). The product development indices are presented in Section 2. The dependent variable is fund family market share, defined as the sum of TNA in the family's active funds divided by the sum of TNA in active funds in the industry. All regressions include lagged controls at the family- and industry-level. Control variables are described in Appendix B.3. Cluster robust standard errors are reported in the parentheses, with the level of cluster presented in the table. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

of fund style. First, I define fund style as the Morningstar categories. However, since the Morningstar categories are highly fractioned it might be more difficult to find valid matches. Therefore, I also run the same matching procedure defining fund style using the Morningstar style-box.[33] Through these matching procedures, the pseudo-families has the same number of funds with the same style, and similar family TNA as the original families.[34] More details on the construction of the pseudo-families are provided in Appendix D.3.

The regressions have the same specifications as the main regression tests of the second hypothesis. Since the families in these samples are made up of funds from other families on a random basis, I do not expect that the results hold in the placebo regressions. First, the results from placebo tests of how product quality development affects fund performance are presented in Table D4. I do not find that product quality development in pseudo-families has a statistically significant effect on performance. Compared to the main results, the coefficient estimates are lower in magnitude. Moreover, standard errors clustered by fund are around the same size, while the independent two-way clustered standard errors are lower in most of the regressions. Thus, the reason why the results are not statistically significant is the magnitude of the coefficients. Next, the results from placebo tests of how product base development affects family market share are presented in Table D5. I find no statistically significant effect of product base development in pseudo-families on net flows or market share. Again, the standard errors are around the same size as in the main results, while the magnitude of the coefficients is smaller.

In sum, my findings in this section show that product quality development makes the family more competitive in terms of performance, and that product base development makes the family more competitive in terms of revenue or market share. These findings support the hypotheses derived in Section 2.3. The findings in the performance hypothesis shows that families can affect the performance of their funds through product development strategies. The results in the market share hypothesis are

[33]The nine styles in the style-box include small-cap value, small-cap blend, small-cap growth, mid-cap value, mid-cap blend, mid-cap growth, large-cap value, large-cap blend, and large-cap growth. See Morningstar Style-Box for more information.
[34]A 100% match of the family is not possible if I am unable to find valid matches for all of the funds in the family.

consistent with the findings of Khorana and Servaes (1999) and Khorana and Servaes (2012), which show that product differentiation is an effective strategy for acquiring market share. Furthermore, product development through the product base channel is not in the favor of the investors, since I find no effect on the performance of the funds. I find some evidence for product quality development being a source of increasing net flows to family-affiliated funds in tests with stricter winsorizing of the net flow variable, but product base development appears to be a better source for increasing the family value. The positive relationship between product quality development and net flows might also be a result of investors chasing past performance (see, e.g., Sirri and Tufano (1998) and Ferreira et al. (2012)). The opposite sign of the coefficient for net quality development index in the tests of the two sub-hypotheses shows that these two channels affect performance and market share differently.

# 6 Discussion and conclusion

Actively managed mutual funds manage trillions of dollars. Therefore, an efficient mutual fund industry is important to ensure investor protection and financial stability. This paper examines whether competition can be a driver of industry efficiency. Despite evidence that fund fee and performance outcomes are affected by competition, little evidence exists of how the incentives of market participants in the mutual fund industry are affected by competition.[35] The level of competition has also been a focus for financial authorities in the last couple of years, and insufficient competition has been identified as a source of inefficiency in the mutual fund industry (see Financial Conduct Authority (2017) and Australian Securities and Investment Commission (2020)).

Feldman et al. (2020) examine alpha-production incentives in light of competition, and find that competition reduces fund managers' willingness to search for alpha. In this paper, I build on this and abstract from performance and fee outcomes as direct causes of competition. I define two channels in which fund families can develop their products in order to stay competitive. The channels are the quality of family-affiliated funds and expansions of the fund base. I focus on the fund families because they are responsible for the strategic decisions relating to their funds and because the family structure in the mutual fund industry can be a source of conflicts of interest (Chevalier and Ellison (1997)).[36]

I highlight the following three results. First, when industry competition increases, fund families respond by increasing product quality development rather than increasing product base development. To account for potential endogeneity issues in the industry competition variables, I extend the tests by running instrumental variable regressions with Bartik-type instruments (Bartik (1991) and Bartik (2002)). In the results, the magnitude of competition on product quality development increases, and thereby confirming that an increase in industry competition is associated with an increase in family product development. I find support for the hypothesis that fund families carry out product development to stay competitive, and that this is done through the quality channel.

Second, fund family product quality development increases the performance of the family-affiliated funds. Third, product base development increases the market share of the fund family. The last two results imply that, when fund families face greater competitive pressure, they choose to increase the quality and performance of their funds, and do not focus on increasing their market share. This is in favor of the investors, and I argue that this is evidence that competition reduces conflicts of interest in the mutual fund industry.

I find no evidence for a reduction in alpha production efforts when competition increases. However, there are two key differences in my analysis compared to Feldman et al. (2020). First, the quality development index does not contain portfolio-level measures. Second, I focus on fund families as opposed to individual funds. Thus, I argue that my results are not in direct contrast to their findings. Moreover, Khorana and Servaes (2012) find that fund families expand their fund base to increase their market share. While I find that this is an efficient channel for increasing the market share, I do not find that it is the primary response to competition.

[35]See, e.g., Coates IV and Hubbard (2007), Gil-Bazo and Ruiz-Verdú (2009), Khorana et al. (2009), Wahal and Wang (2011), and Khorana and Servaes (2012) for studies of fees and Dyck et al. (2013), Pástor et al. (2015), Hoberg et al. (2018), and Leippold and Rueegg (2020) for performance studies.

[36]The source of the conflicts of interest is that the mutual fund family's objective is to maximize the total revenue from its funds, while its investors seek to maximize their risk-adjusted returns.

The scope of this analysis comprises the quality and base channels of product development. The input variables are motivated by the mutual fund literature and include variables that I argue describe the two channels at the family-level. However, these variables do not provide a complete description of all types of product development that take place within these two channels in fund families. Moreover, more general product development, for example product development toward ESG-funds or technological development, is not captured by the product development indices. Together with the growth in popularity of passive index funds, there has also been a major development toward this segment. However, these variables are outside the scope of this paper, and I see this as a possible extension to be addressed in future research. Furthermore, this study relies on industry competition variables that are constructed using standard concentration measures. Another possible extension would be to include additional competition variables, for example company-specific competition variables (see Hoberg et al. (2018) for an example of such a variable). In light of the results and the scope of my analysis, I argue that competition leaves investors better off, and thus reduces conflicts of interest.

## References


Aghion, P., Bechtold, S., Cassar, L. and Herz, H. (2018), 'The causal effects of competition on innovation: Experimental evidence', *The Journal of Law, Economics, and Organization* **34**(2), 162–195.

Amihud, Y. and Goyenko, R. (2013), 'Mutual fund's r 2 as predictor of performance', *The Review of Financial Studies* **26**(3), 667–694.

Australian Securities and Investment Commission (2020), 'ASIC MARKET STUDY CONSULTANCY FUNDS MANAGEMENT INDUSTRY TERMS OF REFERENCE '. Available at https://download.asic.gov.au/media/5727060/terms-of-reference-review-of-competition-in-the-australian-funds-management-industry.pdf, Last accessed on 2021-16-06.

Azar, J., Schmalz, M. C. and Tecu, I. (2018), 'Anticompetitive effects of common ownership', *The Journal of Finance* **73**(4), 1513–1565.

Bartik, T. J. (1991), 'Who benefits from state and local economic development policies?'.

Bartik, T. J. (2002), 'Instrumental variable estimates of the labor market spillover effects of welfare reform', *WE Upjohn Institute Staff Working Paper* (02-78).

Becker, B. and Milbourn, T. (2011), 'How did increased competition affect credit ratings?', *Journal of Financial Economics* **101**(3), 493–514.

Bennett, V. M., Pierce, L., Snyder, J. A. and Toffel, M. W. (2013), 'Customer-driven misconduct: How competition corrupts business practices', *Management Science* **59**(8), 1725–1742.

Berk, J. B. and Green, R. C. (2004), 'Mutual fund flows and performance in rational markets', *Journal of political economy* **112**(6), 1269–1295.

Berk, J. B. and Van Binsbergen, J. H. (2015), 'Measuring skill in the mutual fund industry', *Journal of Financial Economics* **118**(1), 1–20.

Berk, J. B., Van Binsbergen, J. H. and Liu, B. (2017), 'Matching capital and labor', *The Journal of Finance* **72**(6), 2467–2504.

Bessler, W., Blake, D., Lückoff, P. and Tonks, I. (2018), 'Fund flows, manager changes, and performance persistence', *Review of Finance* **22**(5), 1911–1947.

Bryant, L. L. (2012), '"down but not out" mutual fund manager turnover within fund families', *Journal of Financial Intermediation* **21**(4), 569–593.

Bustamante, M. C. and Frésard, L. (2020), 'Does firm investment respond to peers' investment?', *Management Science* .

Chan, C.-Y., Lai, C. W. and Lee, L.-C. (2017), 'Strategic choice of risk: Evidence from mutual fund families', *Journal of Financial Services Research* **51**(1), 125–163.

Chevalier, J. and Ellison, G. (1997), 'Risk taking by mutual funds as a response to incentives', *Journal of political economy* **105**(6), 1167–1200.

Chevalier, J. and Ellison, G. (1999), 'Career concerns of mutual fund managers', *The Quarterly Journal of Economics* **114**(2), 389–432.

Chhaochharia, V., Grinstein, Y., Grullon, G. and Michaely, R. (2017), 'Product market competition and internal governance: Evidence from the sarbanes–oxley act', *Management Science* **63**(5), 1405–1424.

Coates IV, J. C. and Hubbard, R. G. (2007), 'Competition in the mutual fund industry: Evidence and implications for policy', *J. Corp. L.* **33**, 151.

Cornaggia, J., Mao, Y., Tian, X. and Wolfe, B. (2015), 'Does banking competition affect innovation?', *Journal of financial economics* **115**(1), 189–209.

Cremers, K. (2016), 'Do mutual fund investors get what they pay for: Securities law and closet index funds', *Va. L. & Bus. Rev.* **11**, 31.

Cremers, K. M. and Petajisto, A. (2009), 'How active is your fund manager? a new measure that predicts performance', *The review of financial studies* **22**(9), 3329–3365.

Cremers, M., Ferreira, M. A., Matos, P. and Starks, L. (2016), 'Indexing and active fund management: International evidence', *Journal of Financial Economics* **120**(3), 539–560.

Cremers, M., Fulkerson, J. A. and Riley, T. B. (2019), 'Benchmark discrepancies and mutual fund performance evaluation', *Jon A. and Riley, Timothy Brandon, Benchmark Discrepancies and Mutual Fund Performance Evaluation (October 11, 2019)* .

Demirci, I., Ferreira, M. A., Matos, P. and Sialm, C. (2020), How global is your mutual fund? international diversification from multinationals, Technical report, National Bureau of Economic Research.

Dyck, A., Lins, K. V. and Pomorski, L. (2013), 'Does active management pay? new international evidence', *The Review of Asset Pricing Studies* **3**(2), 200–228.

Elton, E. J., Gruber, M. J. and Green, T. C. (2007), 'The impact of mutual fund family membership on investor risk', *Journal of Financial and Quantitative Analysis* **42**(2), 257–277.

Evans, R. B., Prado, M. P. and Zambrana, R. (2020), 'Competition and cooperation in mutual fund families', *Journal of Financial Economics* **136**(1), 168–188.

Fama, E. F. and French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of* .

Fama, E. F. and French, K. R. (1996), 'Multifactor explanations of asset pricing anomalies', *The journal of finance* **51**(1), 55–84.

Fang, J., Kempf, A. and Trapp, M. (2014), 'Fund manager allocation', *Journal of Financial Economics* **111**(3), 661–674.

Feldman, D., Saxena, K. and Xu, J. (2020), 'Is the active fund management industry concentrated enough?', *Journal of Financial Economics* **136**(1), 23–43.

Ferreira, M. A., Keswani, A., Miguel, A. F. and Ramos, S. B. (2012), 'The flow-performance relationship around the world', *Journal of Banking & Finance* **36**(6), 1759–1780.

Ferreira, M. A. and Ramos, S. B. (2009), 'Mutual fund industry competition and concentration: International evidence', *Available at SSRN 1343096* .

Financial Conduct Authority (2017), 'Asset Management Market Study'. Available at https://www.fca.org.uk/publication/market-studies/ms15-2-3.pdf, Last accessed on 2021-16-06.

Gaspar, J.-M., Massa, M. and Matos, P. (2006), 'Favoritism in mutual fund families? evidence on strategic cross-fund subsidization', *The Journal of Finance* **61**(1), 73–104.

Gil-Bazo, J. and Ruiz-Verdú, P. (2009), 'The relation between price and performance in the mutual fund industry', *The Journal of Finance* **64**(5), 2153–2183.

Guedj, I. and Papastaikoudi, J. (2003), Can mutual fund families affect the performance of their funds?, *in* 'EFMA 2004 Basel Meetings Paper'.

Hoberg, G., Kumar, N. and Prabhala, N. (2018), 'Mutual fund competition, managerial skill, and alpha persistence', *The Review of Financial Studies* **31**(5), 1896–1929.

Hortaçsu, A. and Syverson, C. (2004), 'Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds', *The Quarterly journal of economics* **119**(2), 403–456.

Investment Company Institute (2019), 'A Review of Trends and Activities in the Investment Company Industry'. Available at https://ici.org/pdf/2019_factbook.pdf, Last accessed on 2021-21-06.

Jayaraman, N., Khorana, A. and Nelling, E. (2002), 'An analysis of the determinants and shareholder wealth effects of mutual fund mergers', *The Journal of Finance* **57**(3), 1521–1551.

Karuna, C. (2007), 'Industry product market competition and managerial incentives', *Journal of accounting and economics* **43**(2-3), 275–297.

Kempf, A. and Ruenzi, S. (2008), 'Tournaments in mutual-fund families', *The Review of Financial Studies* **21**(2), 1013–1036.

Khorana, A. (1996), 'Top management turnover an empirical investigation of mutual fund managers', *Journal of financial economics* **40**(3), 403–427.

Khorana, A. and Servaes, H. (1999), 'The determinants of mutual fund starts', *The review of financial studies* **12**(5), 1043–1074.

Khorana, A. and Servaes, H. (2012), 'What drives market share in the mutual fund industry?', *Review of Finance* **16**(1), 81–113.

Khorana, A., Servaes, H. and Tufano, P. (2009), 'Mutual fund fees around the world', *The Review of Financial Studies* **22**(3), 1279–1310.

Kostovetsky, L. (2017), 'Brain drain: Are mutual funds losing their best minds?', *Quarterly Journal of Finance* **7**(03), 1750009.

Leippold, M. and Rueegg, R. (2020), 'How rational and competitive is the market for mutual funds?', *Review of Finance* **24**(3), 579–613.

Marxt, C. and Hacklin, F. (2005), 'Design, product development, innovation: all the same in the end? a short discussion on terminology', *Journal of Engineering Design* **16**(4), 413–421.

Massa, M. (2003), 'How do family strategies affect fund performance? when performance-maximization is not the only game in town', *Journal of Financial Economics* **67**(2), 249–304.

Moshirian, F., Tian, X., Zhang, B. and Zhang, W. (2021), 'Stock market liberalization and innovation', *Journal of Financial Economics* **139**(3), 985–1014.

Nanda, V., Wang, Z. J. and Zheng, L. (2004), 'Family values and the star phenomenon: Strategies of mutual fund families', *The Review of Financial Studies* **17**(3), 667–698.

Pástor, L., Stambaugh, R. F. and Taylor, L. A. (2015), 'Scale and skill in active management', *Journal of Financial Economics* **116**(1), 23–45.

Pástor, L., Stambaugh, R. F. and Taylor, L. A. (2017), 'Do funds make more when they trade more?', *The Journal of Finance* **72**(4), 1483–1528.

Pollet, J. M. and Wilson, M. (2008), 'How does size affect mutual fund behavior?', *The Journal of Finance* **63**(6), 2941–2969.

Rodríguez-Castelán, C., López-Calva, L.-F. and Barriga-Cabanillas, O. (2019), 'Market concentration, trade exposure, and productivity in developing countries: Evidence from mexico'.

Schmalz, M. C. (2017), 'Common ownership and competition: Facts, misconceptions, and what to do about it-background paper for oecd hearing on common ownership by institutional investors and its impact on competition'.

Schmidt, K. M., Fey, L. and Thoma, C. (2017), 'Competition and incentives', *European Economic Review* **98**, 111–125.

Sirri, E. R. and Tufano, P. (1998), 'Costly search and mutual fund flows', *The journal of finance* **53**(5), 1589–1622.

Sun, Y. (2020), 'Index fund entry and financial product market competition', *Management Science* .

Tirole, J. and Jean, T. (1988), *The theory of industrial organization*, MIT press.

Ullman, D. G. (1992), *The mechanical design process*, Vol. 2, McGraw-Hill New York.

Ulrich, K. T. (2003), *Product design and development*, Tata McGraw-Hill Education.

Wahal, S. and Wang, A. Y. (2011), 'Competition among mutual funds', *Journal of Financial Economics* **99**(1), 40–59.

Wermers, R. (2003), 'Are mutual fund shareholders compensated for active management "bets"', *Working Study, University of Maryland* .

Zhao, X. (2005), 'Exit decisions in the us mutual fund industry', *The Journal of Business* **78**(4), 1365–1402.

# Appendices

## A Product development indices

### A.1 Details on the construction of product development indices

Here, I provide a detailed example of how the product development indices are constructed for one of the fund families in the sample. The data comes from Morningstar Direct. The data used in the construction of the product quality development index are the Morningstar categories, MPT benchmark-adjusted returns, and data on manager history. To construct the product base development index, I use data on Morningstar categories, inception dates, obsolete dates, and information about liquidation type to detect within-family mergers. I present the details on the raw data and how the components of the indices are constructed for the Vanguard fund family in 2010. For variables defined as the share of existing funds, I also need information on the number of actively managed funds at the end of the previous year.[37] At the end of 2009, the Vanguard family had 31 actively managed funds.

#### Product quality development index

The input variables in the product quality development index are the funds' tracking error, the family's annual manager turnover, within-family performance dispersion, and the family's share of star funds in a year. Tracking error is defined as the average standard deviation of the MPT benchmark-adjusted returns. The tracking error decile rank is computed as the average decile rank of the funds in the family. For the Vanguard family, the average tracking error is equal to 1.27%, while the average decile rank in the funds is close to 4.[38] Manager turnover is defined as the sum of manager changes within the year, divided by the number of funds at the end of the previous year.[39] For the Vanguard fund family in 2010, this value is equal to 7/31.

The performance dispersion is defined as the standard deviation of annual benchmark-adjusted returns across the funds in the family. For the Vanguard family in 2010, this value is equal to 6.40%. The share of star funds is defined as the number of star funds divided by the total number of funds in the family at the end of the previous year. Star funds are defined as funds in the top 5th percentile of MPT benchmark-adjusted returns within a domicile-category. Vanguard had two star funds in 2010, and the value is equal to 2/31.

With the raw family data in hand, I calculate decile ranks for the input variables relative to other families in the entire cross-section (i.e., in the given year). The decile ranks for the product quality development input variables for the Vanguard family in 2010 are as follows:

- Tracking error: 4th decile
- Manager changes: 5th decile
- Performance dispersion: 4th decile
- Star fund: 8th decile

[37] End of year data from Morningstar at the 31st December.
[38] As this is the average decile rank across funds, this decile rank is not necessarily exactly equal to an integer.
[39] For this variable, I divide by the number of funds with data in the field *Manager History*.

The index is defined as the equal-weighted average of the input variable ranks. Thus, the product quality development index for Vanguard family in 2010 is equal to

$$\frac{1}{4}(4+5+4+8)=5.25 \tag{15}$$

**Product base development index**

The input variables in the product base development index include fund starts, fund liquidations, within-family mergers, category starts, and category liquidations. The fund start and liquidation variables do not include funds started in new categories or funds liquidated such that the category is liquidated. They are captured by the category starts and category liquidations variables. The index is defined with respect to fund base expansions, i.e., fund liquidations, category liquidations, and mergers contribute negatively toward the index since these variables do not expand the fund base for the investors.

Starting with funds started in existing categories and funds started in new categories, the Vanguard family started zero new funds in existing categories and one fund in a new category in 2010. As regards mergers, the family carried out zero within-family mergers in 2010. For funds and categories liquidated in 2010, zero funds were liquidated and one fund category was liquidated. All input variables in the product base development index are computed as a share of the number of funds in the family at the end of the previous year.

Like the product quality development index, I calculate decile ranks for the family relative to other families in the entire cross-section within the year. Below are the decile ranks for the product base development input variables for the Vanguard family in 2010:

- Fund starts: 0th decile
- Fund liquidations: 10th decile
- Within-family mergers: 10th decile
- Category starts: 1st decile
- Category liquidations: 9th decile

And the product base development index is equal to

$$\frac{1}{5}(0+10+10+1+9)=6.00 \tag{16}$$

Next, to complete the construction of the product development indices, both of the indices are standardized annually across all funds in the cross-section in each year, with a mean of zero and a standard deviation of one.

## A.2 Summary statistics product development indices

This part of the Appendix presents summary statistics for the product development indices. First, I present the distributions of deviations from within-family means, and next the time series of average deviations from within-family means.

### Distribution of product development indices

Figure A1 presents the distributions of the product development indices demeaned family-by-family, i.e., the variations used in the family-fixed effects regressions in Table 3 and 4. The dependent variables exhibits sufficient variation for meaningful within-family analysis. The distribution of the product quality development is dispersed around zero, on both sides. The distribution of the product base development is skewed toward values below zero. This index will be lower for fund families with no or few fund base expansions during a year. Therefore, if the index is high in some of the years in the sample period, most of the remaining years will be below the average, i.e., a negative difference from the mean. Moreover, the distribution of the total product development and the net product quality development are dispersed around zero.

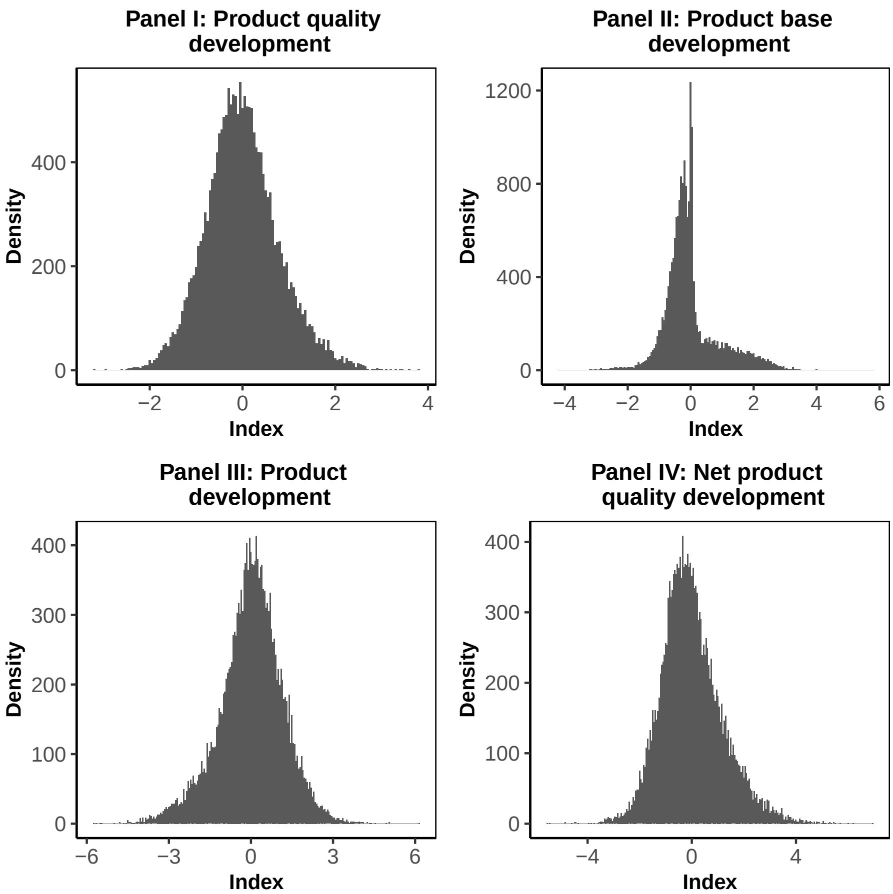



**Figure A1. Product development indices distribution**
Distribution of product development indices. The indices are explained in Section 2.2. All distributions are deviations from the within-family means. Panel I presents the distribution of the product quality development index, Panel II of the product base development index, Panel III of the total product development index, and Panel IV of the net product quality development index.

### Time series of product development indices

The product development indices explained in Section 2.2 are normalized with zero mean each year, which implies that they abstract from time series variation. Therefore, Figure A2 presents the time series of the average deviations from the within-family means. The plot in Panel I is equal-weighted (EW), and the plot in Panel II is value-weighted (VW). Both plots are from 2006 through 2019.

The equal-weighted plot in Panel I shows that there is more product base development in the beginning of the sample period, and that there is relatively more product quality development in the end of the sample period. In the value-weighted plot in Panel II, the development in the two channels are more similar, with more variation from year to year. The difference in the two panels shows that the size of the families matters for how they conduct product development.

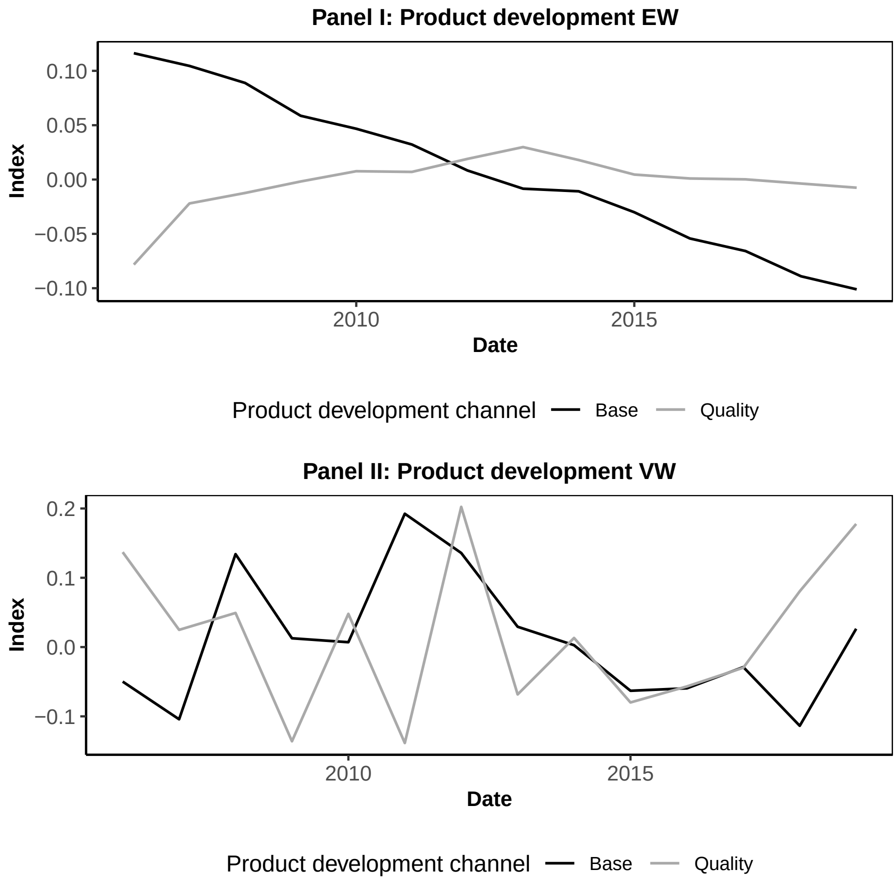



**Figure A2. Product development indices time series**
Development in the product quality and base development indices from 2006 through 2019. The indices are explained in Section 2.2. Both panels plot mean deviations from within-family means, where the families in Panel I are equal-weighted (EW) and the families in Panel II are value-weighted (VW).

# B Description of the data

This appendix presents the steps from downloading the raw data to constructing the final sample. First, the initial base sample is presented. Next, I present details of the construction of the fund and fund family sample. Finally, the control variables from the regressions are presented.

## B.1 Sample selection

The raw sample is constructed using lists from Morningstar based on domiciles. For each domicile, I construct lists consisting of funds categorized as "Open-End Fund" long-only equity funds by Morningstar Direct, including non-survivors from January 2005 through December 2019.[40] The sample includes off-shore funds, typically domiciled in countries such as Ireland, Liechtenstein, and Luxembourg. Countries of domicile are often used to place funds within a country in international mutual fund studies (see, e.g., Cremers (2016) and Demirci et al. (2020)). When splitting the data by domicile, families might have funds located in multiple legal domiciles. Therefore, the same fund family can be present in the data with different domiciles since the families are defined as domicile-family pairs.

The sample is constructed with the aim of including as much of the global sample of actively managed mutual funds as possible. Since the scope of the study comprises funds in which efforts are made in relation to running the funds, funds without a managed portfolio and funds replicating an index are filtered out of the sample. For this, I use the fields: *Fund of Funds*, *Index Fund*, and *Enhanced Index Fund*.[41] These three fields take the values *Yes* or *No*, and I filter such that all funds have the value No. Next, domiciles with less than 50 actively managed funds in the Morningstar database are excluded.[42] After exclusions, I end up with an initial sample of 40,788 actively managed funds from 5,258 fund families. A summary of the initial sample, with the number of funds and fund families in each domicile, is presented in Table B1. The table shows that there are large variations between the different domiciles, measured by the number of families and funds. The US has most families and funds, but Luxembourg also has a relatively high number of funds and families. The smallest industries, measured by the number of families, are Portugal and Chile.

[40] The universe of funds is defined as open-ended funds in Morningstar Direct. The equity funds are categorized by the field *Global Broad Category* equal to equity.

[41] These funds are more homogeneous products, where the potential to differentiate in terms of fund characteristics is limited. Funds of funds are funds limited to investing in other funds, while the index funds' main objective is to track the return of an index.

[42] This is done to impose a lower limit on the competitive pressure.

| Domicile | Number of fund families | Number of funds |
|---|---|---|
| Australia | 278 | 1,767 |
| Austria | 24 | 335 |
| Belgium | 44 | 345 |
| Brazil | 357 | 1,178 |
| Canada | 195 | 2,215 |
| Cayman Islands | 158 | 294 |
| Chile | 23 | 251 |
| China | 86 | 428 |
| Denmark | 73 | 436 |
| Finland | 38 | 354 |
| France | 409 | 2,927 |
| Germany | 65 | 712 |
| Hong Kong | 56 | 192 |
| India | 54 | 462 |
| Indonesia | 64 | 187 |
| Ireland | 462 | 2,154 |
| Israel | 25 | 353 |
| Italy | 56 | 344 |
| Japan | 75 | 3,471 |
| Jersey | 27 | 107 |
| Liechtenstein | 25 | 347 |
| Luxembourg | 755 | 6,612 |
| Malaysia | 41 | 459 |
| Mexico | 26 | 124 |
| Netherlands | 70 | 299 |
| New Zealand | 38 | 199 |
| Norway | 32 | 197 |
| Poland | 27 | 130 |
| Portugal | 18 | 71 |
| Saudi Arabia | 35 | 173 |
| Singapore | 32 | 230 |
| South Africa | 47 | 476 |
| South Korea | 53 | 1,955 |
| Spain | 106 | 632 |
| Sweden | 88 | 497 |
| Switzerland | 68 | 563 |
| Taiwan | 45 | 529 |
| Thailand | 26 | 959 |
| United Kingdom | 276 | 1,913 |
| United States | 881 | 5,911 |
| Sum | 5,258 | 40,788 |

**Table B1. Fund sample**
Summary of all domiciles included in the sample, presenting the total number of fund families and funds. The number of funds and fund families is the number after the initial cleaning of the sample, as explained in Appendix B.1.

### B.2 Fund, family, and industry data

In this section, I explain how the final sample of funds, fund families, and industries are constructed using the initial sample explained in the previous section. Most funds in the sample are structured with multiple share classes. For each share class, I download the following time series variables: gross returns, net returns, net assets, and fund size aggregated over share classes, all with monthly frequency. All of the time series variables are converted into USD in Morningstar Direct. Moreover, I download cross-sectional fund information from the same source. The cross-sectional fund information contains information about the funds' domicile, fund company, Morningstar category, Morningstar style, inception and obsolete date, obsolete type, manager history, as well as other useful information about the funds. I compute expense ratios using gross and net returns following the definition from Morningstar.[43] As reported by Berk and Van Binsbergen (2015), I often observe that net assets are reported quarterly or are missing for a specific month. In this case, I roll the assets under the assumption of zero net flows. I download CPI-data from the data library of the World Bank, and CPI-adjust all variables denoted in USD to the dollar-level as of December 2019.[44]

To obtain the fund-level data from the share class-level data, I aggregate the time series data over share classes. I use the fields *Sec ID* and *Fund ID*, where *Sec ID* is individual for each share class, while *Fund ID* is the unique portfolio ID and is identical for all share classes belonging to the same main fund.[45] I make use of the field *Oldest Share Class* from Morningstar to retrieve the main share class of each fund. The funds are categorized using the field *Morningstar Category* from Morningstar Direct. The benchmark returns used to compute the benchmark-adjusted returns are the MPT Benchmark from Morningstar. This is the primary benchmark assigned to investments based on their Morningstar category, and it is used to calculate Modern Portfolio Theory statistics.[46] In addition, I download regional factor returns from Ken French's website.[47] The funds are manually assigned to a region based on the Morningstar Category. Factor coefficients are obtained by running factor regressions for each fund on the whole sample period of monthly observations. In the next step, I use these factor coefficients together with the factor returns to risk-adjust the fund returns.

The initial data is downloaded with monthly frequency and annualized to obtain fund-year level data. Static variables are collected at year-end. Annual returns are computed as the cumulative monthly returns within the year. Funds with incomplete return data, i.e., less than 12 months of data within the year, are not given an annual return in these years to avoid bias in the data from incomplete fund-year observations.[48] Next, after annualizing the fund data, I construct the main sample of the paper, with annual fund family observations. Then, I aggregate the fund-level data to family-level data, using the fields *Domicile* and *Firm Name ID*.[49] Finally, to construct the annual industry-level data, I use the field *Domicile* in the fund-year and family-year samples. The next section presents the fund-, family-, and industry-level control variables.

[43] Definition available at: Morningstar Gross Return.

[44] Data available at: World Bank Data Library.

[45] These funds have the same portfolios, but differ in their fee structure.

[46] Information about the MPT benchmarks is available at: MPT Benchmarks.

[47] Data available at: Ken French Data Library.

[48] I use the fields *Inception Date* and *Obsolete Date* to filter out funds that has been alive for less than 12 months within a year.

[49] The mutual fund firms are categorized by the fields *Firm Name* and *Firm Name ID*. These firms are what I refer to as fund families throughout the paper.

## B.3 Control variables

This section explains the control variables at different levels of observations. Dependent variables are explained in the main text.

### Fund variables

Expense ratio - Annual expense ratio.

Log total net assets - The natural logarithm of total net assets in the fund (in million USD) CPI-adjusted to the dollar-level in December 2019.

Net flow - Annual sum of monthly net flows as a percentage of total net assets at the end of the previous year.

Log fund age - The natural logarithm of years since the fund inception date.

### Family variables

Log family total net assets - The natural logarithm of total net assets in the family of active funds (in million USD) CPI-adjusted to the dollar-level in December 2019.

Log family age - The natural logarithm of years since the inception date of the family's first fund.

Log number of funds in the family - The natural logarithm of the number of funds in the family.

### Industry variables

Log industry size - The natural logarithm of the sum of total net assets in the industry of open-end equity mutual funds (in million USD) CPI-adjusted to the dollar-level in December 2019.

Log GDP per capita - The natural logarithm of gross domestic product per capita in US dollars. GDP per capita is downloaded from the data library of the World Bank, and is available at: World Bank Data Library.

# C Competition and product development: additional results

This appendix contains additional results from the tests of the first hypothesis.

## C.1 First-difference regressions

I run two types of first-difference regressions. The dependent variables in these regressions are the change in the product development indices from time $t-1$ to $t$ ($\Delta Index$). The main test variables in the regressions are the change in $Comp^{HHI}$ and $Comp^{NHHI}$ from time $t-2$ to $t-1$, defined by $\Delta Comp^{HHI}_{t-1}$ and $\Delta Comp^{NHHI}_{t-1}$. The other control variables are also defined by the difference from $t-2$ to $t-1$.

### Aggregated cross-country tests

First, the regressions are run on a sample aggregated by industry. The dependent variables are the average first-difference from $t-1$ to $t$ in the product development indices. The regressions are OLS regressions described by the specification

$$\Delta Index_{c,t} = a + \gamma \Delta Comp_{c,t-1} + \beta_1 + \beta_1 \Delta X_{c,t-1} + \epsilon_{c,t} \tag{17}$$

where the regressions are run both with and without control variables. The regression results are presented in Table C1.

| | **Panel I: Industry competition** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | ΔProduct quality development | | ΔProduct base development | | ΔProduct development | | ΔNet product quality development | |
| $\Delta Comp^{HHI}_{t-1}$ | 0.861** | 1.104*** | −1.165*** | −1.082*** | −0.304 | 0.021 | 2.026*** | 2.186*** |
| | (0.386) | (0.410) | (0.431) | (0.353) | (0.534) | (0.480) | (0.620) | (0.595) |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Obs | 467 | 440 | 467 | 440 | 467 | 440 | 467 | 440 |
| $R^2$ | 0.015 | 0.030 | 0.026 | 0.021 | 0.001 | 0.004 | 0.041 | 0.049 |
| Adj. $R^2$ | 0.013 | 0.023 | 0.024 | 0.014 | −0.001 | −0.003 | 0.039 | 0.042 |
| | **Panel II: Normalized industry competition** | | | | | | | |
| Dep. var. | ΔProduct quality development | | ΔProduct base development | | ΔProduct development | | ΔNet product quality development | |
| $\Delta Comp^{NHHI}_{t-1}$ | 0.841** | 1.081*** | −1.106*** | −1.025*** | −0.265 | 0.056 | 1.947*** | 2.106*** |
| | (0.365) | (0.388) | (0.412) | (0.335) | (0.502) | (0.446) | (0.595) | (0.571) |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Obs. | 467 | 440 | 467 | 440 | 467 | 440 | 467 | 440 |
| $R^2$ | 0.015 | 0.031 | 0.025 | 0.020 | 0.001 | 0.004 | 0.041 | 0.049 |
| Adj. $R^2$ | 0.013 | 0.024 | 0.023 | 0.014 | −0.001 | −0.003 | 0.039 | 0.042 |

**Table C1. Family product development and competition first-difference regressions aggregated by industry**
Aggregated country-level first-difference regressions testing the effect of competition on fund family product development indices, on the form presented in Equation (17). The product development indices are presented in Section 2. Panel I tests industry competition based on *HHI*, and Panel II tests normalized industry competition based on *NHHI*, both explained in Section 2.1. Regressions with control variables include lagged controls at the industry-level. Control variables are described in Appendix B.3. Robust standard errors are reported in the parentheses. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

### Family-level tests

Second, the first-difference regressions are run on the family-level sample. These first-difference regressions correspond to the panel regressions in Section 4. The regressions are OLS regressions in the following form:

$$\Delta Index_{f,c,t} = a + \gamma \Delta Comp_{c,t-1} + \beta_1 \Delta X_{f,c,t-1} + \beta_2 \Delta X_{c,t-1} + \epsilon_{f,c,t} \tag{18}$$

The results are presented in Table C2.

| **Panel I: Industry competition** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | $\Delta$Product quality development | | $\Delta$Product base development | | $\Delta$Product development | | $\Delta$Net product quality development | |
| $\Delta Comp^{HHI}_{t-1}$ | 0.951*** | 1.491*** | −1.020*** | −0.963** | −0.069 | 0.528 | 1.971*** | 2.454*** |
| | (0.315) | (0.403) | (0.350) | (0.477) | (0.431) | (0.609) | (0.508) | (0.640) |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Obs. | 15,642 | 15,012 | 15,642 | 15,012 | 15,642 | 15,012 | 15,642 | 15,012 |
| $R^2$ | 0.0004 | 0.001 | 0.0004 | 0.061 | 0.00000 | 0.035 | 0.001 | 0.033 |
| Adj. $R^2$ | 0.0004 | 0.001 | 0.0004 | 0.061 | −0.0001 | 0.034 | 0.001 | 0.033 |

| **Panel II: Normalized industry competition** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | $\Delta$Product quality development | | $\Delta$Product base development | | $\Delta$Product development | | $\Delta$Net product quality development | |
| $\Delta Comp^{NHHI}_{t-1}$ | 0.894*** | 1.399*** | −0.959*** | −0.966** | −0.065 | 0.433 | 1.853*** | 2.365*** |
| | (0.302) | (0.388) | (0.336) | (0.458) | (0.412) | (0.584) | (0.488) | (0.615) |
| Controls | No | Yes | No | Yes | No | Yes | No | Yes |
| Obs. | 15,642 | 15,012 | 15,642 | 15,012 | 15,642 | 15,012 | 15,642 | 15,012 |
| $R^2$ | 0.0004 | 0.001 | 0.0004 | 0.061 | 0.00000 | 0.035 | 0.001 | 0.033 |
| Adj. $R^2$ | 0.0004 | 0.001 | 0.0003 | 0.061 | −0.0001 | 0.034 | 0.001 | 0.033 |

**Table C2. Family product development and competition first-difference regressions**
Family-level first-difference regressions testing the effect of competition on fund family product development indices, on the form presented in Equation (18). The product development indices are presented in Section 2. Panel I tests industry competition based on *HHI*, and Panel II tests normalized industry competition based on *NHHI*, both explained in Section 2.1. Regressions with control variables include lagged controls at the family- and industry-level. Control variables are described in Appendix B.3. Robust standard errors are reported in the parentheses. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

## C.2 Sample ex-US

As described in the main text, I run the tests of product development and competition on a sample excluding the US active fund industry. The regressions are presented in Table C3.

| **Panel I: Industry competition** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Dep. var. | Product quality development | | Product base development | | Product development | | Net product quality development | |
| $Comp^{HHI}_{t-1}$ | 0.589** | 0.589* | −0.850** | −0.850** | −0.261 | −0.261 | 1.438*** | 1.438** |
| | (0.273) | (0.274) | (0.402) | (0.363) | (0.457) | (0.356) | (0.512) | (0.536) |
| Controls | Family and Domicile | | | | | | | |
| FE | Family + Year | | | | | | | |
| Cluster | Family | Family + Year | Family | Family + Year | Family | Family + Year | Family | Family + Year |
| Obs. | 13,416 | 13,416 | 13,416 | 13,416 | 13,416 | 13,416 | 13,416 | 13,416 |
| $R^2$ | 0.374 | 0.374 | 0.274 | 0.274 | 0.318 | 0.318 | 0.328 | 0.328 |
| Adj. $R^2$ | 0.292 | 0.292 | 0.180 | 0.180 | 0.229 | 0.229 | 0.240 | 0.240 |
| **Panel II: Normalized industry competition** | | | | | | | | |
| Dep. var. | Product quality development | | Product base development | | Product development | | Net product quality development | |
| $Comp^{NHHI}_{t-1}$ | 0.609** | 0.609** | −0.751* | −0.751* | −0.142 | −0.142 | 1.360*** | 1.360** |
| | (0.261) | (0.253) | (0.389) | (0.350) | (0.442) | (0.341) | (0.494) | (0.507) |
| Controls | Family and Domicile | | | | | | | |
| FE | Family + Year | | | | | | | |
| Cluster | Family | Family + Year | Family | Family + Year | Family | Family + Year | Family | Family + Year |
| Obs. | 13,416 | 13,416 | 13,416 | 13,416 | 13,416 | 13,416 | 13,416 | 13,416 |
| $R^2$ | 0.374 | 0.374 | 0.274 | 0.274 | 0.318 | 0.318 | 0.328 | 0.328 |
| Adj. $R^2$ | 0.292 | 0.292 | 0.180 | 0.180 | 0.229 | 0.229 | 0.240 | 0.240 |

**Table C3. Family product development and competition regressions ex-US**
Regressions testing the effect of competition on fund family product development indices, on the form presented in Equation (9), on a sample excluding the US. The product development indices are presented in Section 2. Panel I tests industry competition based on *HHI*, and Panel II tests normalized industry competition based on *NHHI*, both explained in Section 2.1. All regressions include lagged controls at the family- and industry-level. Control variables are described in Appendix B.3. Cluster robust standard errors are reported in the parentheses, with the level of clusters reported in the table. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

### C.3 Bartik instrument details

This part of the appendix explains how the Bartik instruments for the industry competition variables are constructed. The Bartik instrument of variable $y$ is constructed used the starting point of $y$ at time $t = 0$ in country $c$ and the external growth of $y$. The starting point of the industry concentration variables are December 2005. To construct the instrument using the US growth rate, I compute for each country and each year

$$y_{c,t}^{IV} = y_{c,t=2005} \cdot g_t^{US} \tag{19}$$

where $y_{c,t=2005}$ is the industry concentration in country $c$ in December 2005, i.e., the first year of the data, and $g_t^{US}$ represents the US growth rate in industry concentration between period $t = 0$ and $t$. For global growth, I construct the instrument in the same way, but here the growth rate is the value-weighted global growth rate.

$$y_{c,t}^{IV} = y_{c,t=2005} \cdot g_t^{Global} \tag{20}$$

The instrumented variables are the industry concentration variables Herfindahl-Hirschman index ($HHI$) and the normalized Herfindahl-Hirschman index ($NHHI$). To construct the industry competition variables, I subtract the industry concentration variables from 1, as explained in Section 2.1.

# D Product development outcomes: additional results

This appendix contains additional results from the tests of the second hypothesis of product development outcomes.

## D.1 Value added

First, I explain how the annual value added variable is constructed. The monthly value added is computed by

$$VA_t = \alpha_t TNA_{t-1} \tag{21}$$

where $TNA_{t-1}$ is the total net assets at the fund in end of the previous month. I compute the value added using the four different alpha estimates that I run in the main regression in Table 6. Annual value added is computed as the sum of monthly value added within a year. The regression results from using value added estimates as dependent variables are presented in Table D1.

| | Panel I: MPT benchmark adjusted value added | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Index | Product quality development | | Product base development | | Product development | | Net product quality development | |
| $Index_{t-1}$ | 0.741*** | 0.741 | 0.131 | 0.131 | 0.408*** | 0.408 | 0.270** | 0.270 |
| | (0.174) | (0.423) | (0.161) | (0.172) | (0.118) | (0.233) | (0.118) | (0.208) |
| Controls | Fund + Family + Domicile | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 147,659 | 147,659 | 147,659 | 147,659 | 147,659 | 147,659 | 147,659 | 147,659 |
| $R^2$ | 0.163 | 0.163 | 0.162 | 0.162 | 0.163 | 0.163 | 0.162 | 0.162 |
| Adj. $R^2$ | 0.041 | 0.041 | 0.041 | 0.041 | 0.041 | 0.041 | 0.041 | 0.041 |
| | **Panel II: CAPM adjusted value added** | | | | | | | |
| Index | Product quality development | | Product base development | | Product development | | Net product quality development | |
| $Index_{t-1}$ | 0.202 | 0.202 | 0.035 | 0.035 | 0.110 | 0.110 | 0.073 | 0.073 |
| | (0.197) | (0.580) | (0.179) | (0.252) | (0.134) | (0.339) | (0.130) | (0.242) |
| Controls | Fund + Family + Domicile | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 |
| $R^2$ | 0.163 | 0.163 | 0.163 | 0.163 | 0.163 | 0.163 | 0.163 | 0.163 |
| Adj. $R^2$ | 0.042 | 0.042 | 0.042 | 0.042 | 0.042 | 0.042 | 0.042 | 0.042 |
| | **Panel III: FF3 adjusted value added** | | | | | | | |
| Index | Product quality development | | Product base development | | Product development | | Net product quality development | |
| $Index_{t-1}$ | 0.188 | 0.188 | −0.132 | −0.132 | 0.013 | 0.013 | 0.158 | 0.158 |
| | (0.172) | (0.372) | (0.154) | (0.255) | (0.117) | (0.236) | (0.113) | (0.195) |
| Controls | Fund + Family + Domicile | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 |
| $R^2$ | 0.153 | 0.153 | 0.153 | 0.153 | 0.153 | 0.153 | 0.153 | 0.153 |
| Adj. $R^2$ | 0.031 | 0.031 | 0.031 | 0.031 | 0.031 | 0.031 | 0.031 | 0.031 |
| | **Panel IV: FF3MOM adjusted value added** | | | | | | | |
| Index | Product quality development | | Product base development | | Product development | | Net product quality development | |
| $Index_{t-1}$ | 0.171 | 0.171 | −0.134 | −0.134 | 0.004 | 0.004 | 0.151 | 0.151 |
| | (0.171) | (0.459) | (0.152) | (0.233) | (0.115) | (0.271) | (0.111) | (0.209) |
| Controls | Fund + Family + Domicile | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 | 136,851 |
| $R^2$ | 0.156 | 0.156 | 0.156 | 0.156 | 0.156 | 0.156 | 0.156 | 0.156 |
| Adj. $R^2$ | 0.034 | 0.034 | 0.034 | 0.034 | 0.034 | 0.034 | 0.034 | 0.034 |

**Table D1. Family product development and fund value added regressions**

Regressions testing the effect of fund family product development indices on fund value added, on the form presented in Equation (10). The product development indices are presented in Section 2. The dependent variables in the regressions in Panel I is MPT-adjusted value added, dependent variable in Panel II is CAPM-adjusted value added, dependent variable in Panel III is FF3-adjusted value added, and dependent variable in Panel IV is FF3MOM-adjusted value added. All regressions include lagged controls at the fund-, family-, and industry-level. Control variables are described in Appendix B.3. Cluster robust standard errors are reported in the parentheses, with the level of cluster presented in the table. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

## D.2 Net flow regressions

The summary statistics in Table 1 show that there are some outliers at the top of the distribution in fund net flows. As a robustness test to examine whether the extreme values in net flows drive the results, I run the same regressions as in Table D2, with net flows winsorized at the bottom and top 3% and 5% level of the distribution within the whole sample. The regression results are presented in Table D2.

| Panel I: Net flow (%) winsorized at 3% | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Index | Product quality development | | Product base development | | Product development | | Net product quality development | |
| $Index_{t-1}$ | 0.333* | 0.333 | 0.433*** | 0.433 | 0.385*** | 0.385** | −0.082 | −0.082 |
| | (0.172) | (0.225) | (0.163) | (0.243) | (0.118) | (0.147) | (0.119) | (0.184) |
| Controls | Fund + Family + Industry | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 |
| $R^2$ | 0.356 | 0.356 | 0.356 | 0.356 | 0.356 | 0.356 | 0.356 | 0.356 |
| Adj. $R^2$ | 0.263 | 0.263 | 0.263 | 0.263 | 0.263 | 0.263 | 0.263 | 0.263 |
| **Panel II: Net flow (%) winsorized at 5%** | | | | | | | | |
| Index | Product quality development | | Product base development | | Product development | | Net product quality development | |
| $Index_{t-1}$ | 0.317** | 0.317 | 0.317** | 0.317* | 0.315*** | 0.315** | −0.027 | −0.027 |
| | (0.132) | (0.178) | (0.125) | (0.171) | (0.090) | (0.107) | (0.091) | (0.138) |
| Controls | Fund + Family + Industry | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 | 148,932 |
| $R^2$ | 0.364 | 0.364 | 0.364 | 0.364 | 0.364 | 0.364 | 0.364 | 0.364 |
| Adj. $R^2$ | 0.272 | 0.272 | 0.273 | 0.273 | 0.273 | 0.273 | 0.272 | 0.272 |

**Table D2. Family product development and net flow robustness regressions**
Regressions testing the effect of fund family product development indices on fund net flow, on the form presented in Equation (10). The product development indices are presented in Section 2. The dependent variable is net flow in percentage of TNA, where net flows are winsorized at the top and bottom 3% level in Panel I and at the top and bottom 5% level in Panel II. All regressions include lagged controls at the fund-, family-, and industry-level. Control variables are described in Appendix B.3. Cluster robust standard errors are reported in the parentheses, with the level of cluster presented in the table. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

## D.3 Placebo tests of hypothesis 2

This part of the appendix presents details on the construction of the pseudo-families and the placebo tests of the second hypothesis.

### Construction of pseudo-families

First, I explain how the pseudo-families are constructed. Each fund in a family is matched to a fund from another family based on domicile, fund style, and the decile rank of fund size within the domicile. The fund size decile ranks are formed each year, within each domicile. The matching is one-to-one, i.e., a fund is only matched to one other fund from another family, to avoid bias in the pseudo-families from funds being a part of multiple families. In this matching procedure, the pseudo-families consists of funds from the same domicile as the original families, similar number of funds, similar composition of funds based on fund styles, and the families are roughly the same size. I run two different matching procedures, where the difference is the definition of fund style. First, I define fund style as the Morningstar categories. Then, the funds are matched to another fund from the same domicile, with the same Morningstar category, and from the same size decile. Since the Morningstar categories are highly fractioned, it might be more difficult to find valid matches. Therefore, I also run the same matching procedure using fund style defined by the Morningstar style-box.[50] In this matching procedure, the funds are matched to another fund from the same domicile, with the same Morningstar style, and from the same size decile.

Table D3 presents summary statistics for the pseudo-families from the two matching procedures, and comparison with the original families. I am able to match around 70% of the sample in both of the matching procedures, measured in TNA. The reason why it is not a 100% match is that the matching is one-to-one and requires that funds are in the same Morningstar category or style within the domiciles, as well as being in the same fund size decile. In the case of multiple valid matches, the fund with most similar TNA is used as match. However, for the funds that are matched, the families consist of similar funds in terms of size and style composition.

| **Families** | **Original** | **Matched on category** | | **Matched on style** | |
|---|---|---|---|---|---|
| | | Matched | Share | Matched | Share |
| Family TNA (million USD) | 6,240 | 4,298 | 0.69 | 4,396 | 0.70 |
| Number of categories/styles | 6.89 | 4.79 | 0.69 | 5.14 | 0.75 |
| Number of funds | 11.66 | 7.48 | 0.64 | 7.63 | 0.65 |

**Table D3. Family characteristics in matched pseudo-families**
Characteristics of original families and psuedo-families. The pseudo-families are matched with two different matching procedures, both based on domicile and TNA segments, with difference in how fund style is defined. First it is defined as Morningstar categories, and next as Morningstar styles from the style-box.

If the results in the main tests of the second hypothesis are driven by the family composition, number of funds, or family size, the results will also hold in the placebo tests. If the results are in fact driven by coordinated fund family product development, the results in the placebo tests should not be statistically significant. In the next two sections, I run the tests of the second hypothesis on the two samples of

[50] The nine styles in the style-box include small-cap value, small-cap blend, small-cap growth, mid-cap value, mid-cap blend, mid-cap growth, large-cap value, large-cap blend, and large-cap growth. See Morningstar Style-Box for more information.

pseudo-families, matched according to the abovementioned matching procedures. The performance placebo regressions are testing the product quality development index, and the market share placebo regressions are testing the product base development index.

### Performance and product quality development placebo tests

Here, I present the placebo tests of product quality development on the performance of family-affiliated funds. The pseudo-families are constructed based on the methodology explained above. Table D4 presents the placebo tests, with families matched on Morningstar categories in Panel I, and families matched on Morningstar styles in Panel II.

| | **Panel I: Matched based on Morningstar category** | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Dependent variable | MPT adjusted returns | | CAPM adjusted returns | | FF3 adjusted returns | | FF3MOM adjusted returns | |
| $Quality_{t-1}$ | 0.045 | 0.045 | −0.036 | −0.036 | −0.062 | −0.062 | −0.060 | −0.060 |
| | (0.041) | (0.048) | (0.048) | (0.063) | (0.043) | (0.058) | (0.043) | (0.064) |
| Controls | Fund + Family + Industry | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 80,541 | 80,541 | 76,578 | 76,578 | 76,578 | 76,578 | 76,578 | 76,578 |
| $R^2$ | 0.253 | 0.253 | 0.262 | 0.262 | 0.259 | 0.259 | 0.255 | 0.255 |
| Adj. $R^2$ | 0.075 | 0.075 | 0.087 | 0.087 | 0.084 | 0.084 | 0.078 | 0.078 |
| | **Panel II: Matched based on Morningstar style** | | | | | | | |
| Dependent variable | MPT adjusted returns | | CAPM adjusted returns | | FF3 adjusted returns | | FF3MOM adjusted returns | |
| $Quality_{t-1}$ | 0.052 | 0.052 | 0.014 | 0.014 | −0.047 | −0.047 | −0.024 | −0.024 |
| | (0.039) | (0.055) | (0.045) | (0.053) | (0.041) | (0.058) | (0.040) | (0.063) |
| Controls | Fund + Family + Industry | | | | | | | |
| FE | Fund + Year | | | | | | | |
| Cluster | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year | Fund | Fund + Year |
| Obs. | 85,599 | 85,599 | 79,986 | 79,986 | 79,986 | 79,986 | 79,986 | 79,986 |
| $R^2$ | 0.243 | 0.243 | 0.258 | 0.258 | 0.253 | 0.253 | 0.248 | 0.248 |
| Adj. $R^2$ | 0.067 | 0.067 | 0.076 | 0.076 | 0.070 | 0.070 | 0.063 | 0.063 |

**Table D4. Family product quality development and performance placebo regressions**
Regressions testing the effect of pseudo-family product quality development on fund performance, on the form presented in Equation (10). The construction of pseudo-families are described in Appendix D.3. The product quality development index is presented in Section 2. The regressions in Panel I are run on a sample of pseudo-families matched based on domicile, Morningstar category, and fund size. The regressions in Panel II are run on a sample of pseudo-families matched based on domicile, Morningstar style-box, and fund size. Dependent variables are denoted in the table. All regressions include lagged controls at the fund-, family-, and industry-level, as well as lagged adjusted returns, adjusted in the same way as the dependent variable. Control variables are described in Appendix B.3. Cluster robust standard errors are reported in the parentheses, with the level of cluster presented in the table. Asterisks denote statistical significance: $^{***}$p<0.01, $^{**}$p<0.05, $^{*}$p<0.1.

### Market share and product base development placebo tests

Here, I present the placebo tests of product base development on the market share of the family. The pseudo-families are constructed based on the methodology explained above. Table D5 presents the placebo tests, with families matched on Morningstar categories in Panel I, and families matched on Morningstar styles in Panel II.

| | **Panel I: Matched based on Morningstar category** | | | |
|---|---|---|---|---|
| Dependent variable | Fund net flow (%) | | Pseudo-family market share (%) | |
| $Base_{t-1}$ | −0.005 | −0.005 | −0.004 | −0.023 |
| | (0.412) | (0.520) | (0.013) | (0.014) |
| Controls | Fund + Family + Industry | | No | Family + Industry |
| FE | Fund + Year | | Family + Year | |
| Cluster | Fund | Fund + Year | Family | Family + Year |
| Obs. | 80,851 | 80,851 | 13,890 | 13,517 |
| $R^2$ | 0.412 | 0.412 | 0.839 | 0.845 |
| Adj. $R^2$ | 0.272 | 0.272 | 0.817 | 0.824 |
| | **Panel II: Matched based on Morningstar style** | | | |
| Dependent variable | Fund net flow (%) | | Pseudo-family market share (%) | |
| $Base_{t-1}$ | 0.189 | 0.189 | −0.001 | −0.009 |
| | (0.376) | (0.338) | (0.012) | (0.012) |
| Controls | Fund + Family + Industry | | No | Family + Industry |
| FE | Fund + Year | | Family + Year | |
| Cluster | Fund | Fund + Year | Family | Family + Year |
| Obs. | 85,856 | 85,856 | 14,142 | 13,732 |
| $R^2$ | 0.408 | 0.408 | 0.855 | 0.861 |
| Adj. $R^2$ | 0.272 | 0.272 | 0.836 | 0.842 |

**Table D5. Family product base development and market share placebo regressions**
Placebo regressions testing the effect of pseudo-family product base development on fund net flow and family market share, on the form presented in Equation (10) and (11). The construction of pseudo-families are described in Appendix D.3. The product base development index is presented in Section 2. The regressions in Panel I are run on a sample of pseudo-families matched based on domicile, Morningstar category, and fund size. The regressions in Panel II are run on a sample of pseudo-families matched based on domicile, Morningstar style-box, and fund size. Dependent variables are denoted in the table. The fund-level net flow regressions include lagged controls at the fund-, family-, and industry-level. The family-level market share regressions include lagged controls at the family- and industry-level. Control variables are described in Appendix B.3. Cluster robust standard errors are reported in the parentheses, with the level of cluster presented in the table. Asterisks denote statistical significance: $^{***}$p<0.01, $^{**}$p<0.05, $^{*}$p<0.1.

# Forced to be active: Evidence from a regulation intervention*

Petter Bjerksund
Norwegian School of Economics

Trond Døskeland
Norwegian School of Economics

André Wattø Sjuve
Norwegian School of Economics

Andreas Ørpetveit
Norwegian School of Economics

**Abstract**

Closet indexers are low-activity mutual funds sold and marketed as active. Their investors are thus only partially receiving the service they pay for. Therefore, regulation is considered by supervisory authorities worldwide. We examine the impact of policy scrutiny by comparing scrutinized closet index funds in Scandinavia with similar unaffected European funds. We find that funds under scrutiny choose to increase activity over reducing fees and updating their investor information to reflect realized strategy. Despite investors getting a more actively managed fund, our findings suggest that value creation decreases. Regulation thus results in the worst of all worlds.

JEL Classification: D14; G11; K12

Keywords: Asset management; Active portfolio management; Regulation of financial services

*Bjerksund and Døskeland have been expert witnesses for the Norwegian Consumer Council in the lawsuit described in the paper between the Consumer Council and DNB.

# 1 Introduction

Actively managed mutual funds sell the potential to beat their benchmark (usually a market index). Investors who choose this type of fund are typically looking for an opportunity to outperform the market. They pay a premium over index funds for the service of dedicated fund managers who try to beat the market.[1] The fund managers' effort may justify this extra cost if it creates an opportunity to make excess returns by deviating from the fund's benchmark. However, several papers have identified funds with relatively high fees, which, at the same time, have a low degree of active management (see e.g., Cremers and Petajisto (2009), Petajisto (2013), or Cremers, Ferreira, Matos and Starks (2016)). Thus, investors pay for an active portfolio management service without receiving it. These funds are labeled "closet indexers".

It is not only in the academic world that one has addressed the problem of closet index funds; it is also relevant for financial supervisory authorities around the world.[2] The Scandinavian countries were early to put pressure on potential closet indexers.[3] The supervisory authorities held extensive investigations. The recommendations directed at closet indexers were either to update investor information and reduce the fee or increase the fund's activity. This paper examines the impact of policy scrutiny by comparing Scandinavian closet index funds under scrutiny with unaffected European closet index funds.

We hypothesize that funds under scrutiny will choose to increase activity and maintain their fees. Cutting fees is expensive, but increasing activity is comparably cheap. Rephrased, the funds "force" themselves to increase activity. Furthermore, we investigate the consequences of increased activity for investors. One constraint discussed prominently in the active management literature is decreasing returns to scale, i.e., that the quality of marginal investment opportunities declines as active assets under management (AUM) increases. Under no anti-closet-indexing constraint, managers will choose a (subjective) threshold alpha and closet-index assets when opportunities fall below that point. Gross investor returns are maximized when that threshold alpha is zero (Berk and Green, 2004). However, managers might use different threshold alphas. If they are risk or effort averse, they will choose a positive threshold alpha. If they are overconfident, risk-seeking (e.g., due to convex incentives), or desire to signal skill via activity, they will choose a negative one.[4]

Using regulation to induce additional active management can benefit investors if the threshold is positive but not if it is negative. Our empirical work speaks to this by analyzing the effect on investor returns of inducing additional active management. If managers have already exploited their good investment ideas, then more activity will lead to lower excess returns. Conversely, a positive threshold alpha implies that investors will benefit from an increased activity level.

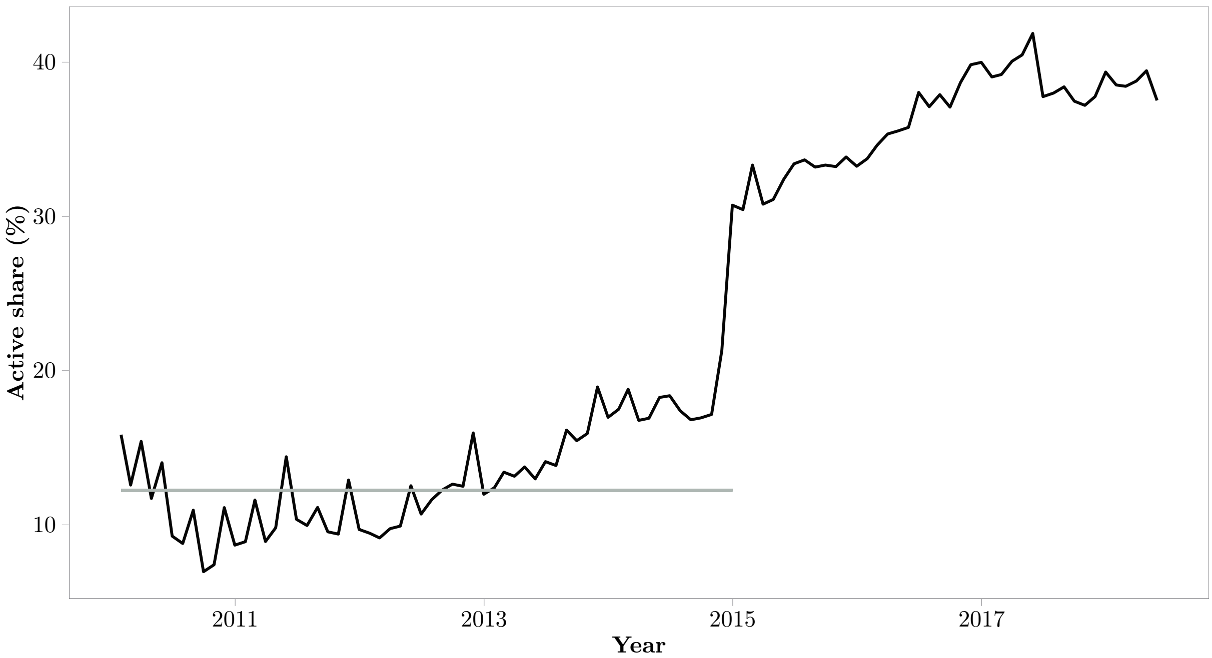
Before we report our findings, we begin with the story of closet indexing in Scandinavia, placing a particular focus on the first closet index fund in the world that was convicted and forced to repay its investors.[5] Closet indexing has been a concern in all Scandinavian countries. The region's different

[1] See for example Morningstar (2019).
[2] See, for example, ESMA (2016), New York Office of the Attorney General (2018), Authority (2019), and ESMA (2020).
[3] See, for example, Reuters article (2017) or Kjørven (2019).
[4] What is optimal for the investor is not necessarily optimal for the fund manager. There are many reasons for frictions. For example, the investor and the manager might have different skill or information sets (see e.g., Sirri and Tufano (1998) and Agarwal, Gay and Ling (2014)), risk preferences (see e.g., Huang, Sialm and Zhang (2011)), or incentives (see e.g., Sensoy (2009) and Ferreira, Matos and Pires (2018)).
[5] There are some cases where the asset managers have accepted a fine (Portfolio Adviser article (2019)) or where there are ongoing lawsuits (Advisor's Edge article (2019)).

national-level Financial Supervisory Authorities (FSAs) have identified questionable practices related to closet indexing (Kjørven, 2019). As a special case, we describe the Norwegian mutual fund, DNB Norge, managed by the largest banking group in Norway, DNB. Investor information and an annual fee of 1.8% suggested that investors should expect active management. Figure 1 plots the development of active share for DNB Norge from the start of 2010 until May 2018. The period the Norwegian FSA investigated DNB Norge was from 2010 until 2014. A quantitative analysis of the degree of active management identified a tracking error at 1.28% per year and an average active share at 12.25%. Based on the discrepancy between what the investors had reason to expect and the actual degree of active management, the FSA concluded that DNB Norge was a closet index fund.

**Figure 1.**
**Active share for DNB Norge**

This figure presents the progression of active share for the mutual fund DNB Norge over the period 2010-2018. The horizontal line at 12.25% between 2010 and 2014 is the average active share for the fund leading up to the corrective order issued by the Norwegian financial supervisory authority.




The Norwegian FSA imposed a corrective order on DNB and gave the fund two alternatives.[6] The first one was to bring the fund's management in line with active management characteristics, as reflected by the fee and the fund's prospectus. Alternatively, the fund could adjust the fee to a level that reflected the applied investment strategy. DNB decided to do a combination of the two. They lowered the fee from 1.8% to 1.4% and more than doubled the active share. From Figure 1 we see the increase at the end of 2014.

Based on the corrective order, the Norwegian Consumer Council (NCC) filed a class-action lawsuit on behalf of the fund's 180,000 investors.[7] The Consumer Council claimed that DNB had misled and overcharged its investors with excessive management fees. In the first instance, the Oslo District Court, DNB won. In the second instance, the Borgarting Court of Appeal ruled in favor of the consumers.

[6] This corrective order is available with an English translation at Finanstilsynet.
[7] For details see Norwegian Consumer Council (2015), and Kjørven (2019).

DNB was ordered to pay back 0.8% of the management fee. The fund then appealed to the Norwegian Supreme Court. The ruling was delivered on the 27th of February 2020, upholding the Court of Appeal's judgment. The verdict can be found here. In Appendix A, we provide more details about the lawsuit. The DNB case is unique in the sense that it has brought the closet index problem to court.[8] However, as we show in this paper, the corrective order on DNB and other similar Scandinavian examples have had spill-over effects on potential index huggers. Many funds changed behavior to avoid being classified as a closet index fund. The consequences of this new behavior is largely unknown. Closet index funds have become more active, but is this also in the interest of the investors?

At one level, the impact of the intervention has an intrinsic value. The natural approach is to interpret the difference between before and after as a result of the intervention. On the other hand, we can never be sure whether the effects are due to the interference or, for example, a general change in the conditions for active management. It could be that at the same time as the attention of the FSAs, there was also troubling times for active management. Then we cannot be certain of what causes a potential low value creation in active management. These endogeneity concerns must be addressed.

To identify the impact of the intervention, we use the exogenous variation created by the policy scrutiny. The estimates of the scrutiny effect come from a difference-in-differences (diff-in-diff) research design, where we compare outcomes from Scandinavian closet index funds under scrutiny (treated funds) with European closet index funds not under scrutiny (control funds). We also use a control group with a selection bias problem in robustness tests, namely truly active Scandinavian funds. It is reassuring for the robustness of the effects that results from both these groups are mostly consistent.

We split the impact assessment into two steps. First, we investigate the impact on the behavior of active management of the funds before we, in the second step, examine the consequences on investors' returns. Regarding the first step, we find the non-surprising result that the fund managers choose to increase activity over updating investor information according to realized strategy. Regarding the second step, we find that the closet index funds that become more active perform worse than a comparable control group of closet index funds. We find that the treated funds put under scrutiny had, on average, an annual active return that was 70 basis points lower than comparable funds unaffected by scrutiny. Based on this thorough examination, we find that going after closet index funds will produce unwillingly active funds, which may be the worst of all worlds. The best solution is not to force these funds to become more active, but to promote fee reduction. If the regulatory authorities had followed the alternative with updating investor information and reducing fees, the investors would fare even better. Assuming that fees had been reduced by half of what DNB had to pay back, the annual increase in value creation would be $1.1\%(= 0.7\% + 0.4\%)$ higher.

The rest of the paper is structured as follows. In Section 2, we discuss related literature and emphasize our relative contribution. To assess the intervention's effect, we use a framework proposed by the UK Financial Conduct Authority (Financial Conduct Authority, 2016). First, in Section 3: Problem diagnosis, we develop an understanding of the problems with closet indexing. In Section 4: Intervention design, we describe the intervention, and present the data with summary statistics. In Section 5: Impact assessment, we examine how fund managers respond to the scrutiny and how it affected the investors. In Section 6, we go into more details and examine the robustness of our previous

[8]There is a new class-action coming up in Canada. For more information, see Newswire article (2020).

findings. We have tried to keep the number of analyzes in the main text to a reasonable level, and we refer to the appendices for additional material. Finally, we suggest some policy implications and conclude in Section 7.

## 2 Related literature and contribution

There is an extensive literature on closet indexing. Even if Berk and Green (2004) mention the term in their paper, it was not until the introduction of the Active share measure by Cremers and Petajisto (2009), that the problem became a subject of detailed study and media attention. Many papers have documented that closet indexing exist all over the world (Cremers et al., 2016).[9] However, there is a difference between identifying the problem and solving it. With market failure, regulation is often called for, but it is not always the best solution. Regulation may make matters worse. The literature about regulating closet index funds is small; one exception is Cremers and Quinn (2016). Least of all is the literature examining the impact of regulating these funds. Hopefully, our assessment can be of interest to other markets (such as the U.S. mutual fund market).

The intervention also provides insight into a more general question; whether regulatory agencies should intervene to correct market failures. The increased complexity of financial arrangements poses a challenge to households managing their financial affairs, and regulators attempting to assist them. There has long been a tension in economics between laissez-faire economists who appreciate and defend the performance of free markets and interventionists who identify market failures and argue that feasible policies can be found to correct them (Campbell, 2016). When households cannot manage their financial decisions, they make mistakes that lowers their welfare and that can have broader consequences for the economy. However, regulation can fail and instead increase the problem it set out to solve. Generally, it is hard to perform cost-benefit analysis in financial regulation (Sunstein, 2015). Our study provides insight into how to assess the effects of the intervention.[10]

Furthermore, we contribute to the literature on the scalability of active management. The extent to which an active fund can outperform its passive benchmark depends not only on the fund's raw skill in identifying investment opportunities, but also on various constraints the fund faces. One constraint that is prominently discussed in recent literature is decreasing returns to scale. If scale impacts performance, skill and scale interact: For example, a larger and more skilled fund can underperform a less skilled small fund. Therefore, to identify the skill, we must also understand the effects of scale.

Berk and Green (2004)'s theoretical model relies on the key assumption of decreasing returns to scale because managers eventually run out of ideas and cannot generate additional alpha. Yet, the empirical evidence on the relationship between fund size and performance is mixed. For U.S. funds, at the fund-level, Chen, Hong, Huang and Kubik (2004), Edelen, Evans and Kadlec (2007), and Yan (2008) document a significantly negative relation between size and performance. However, these findings are challenged by studies that point out an endogeneity issue in the test of the return-to-scale property. Reuter and Zitzewitz (2010) examine the size-performance relation in a natural experiment

[9] Most of the literature focus on the connection between active share and value creation, see for example, Cremers and Petajisto (2009), Hitesh Doshi and Simutin (2015) and ESMA (2020).

[10] Financial Conduct Authority (2016) gives an overview of the British approach to an economic analysis of financial regulation. They outline a methodology for regulatory economic analysis that contemplates a three-stage process, including problem diagnosis, intervention design, and impact assessment.

setting, applying a regression discontinuity approach. Pástor, Stambaugh and Taylor (2015) and Zhu (2018) address the omitted-variable bias by including fund fixed effects to account for heterogeneity in managerial skills. Phillips, Pukthuanthong and Rau (2018) use instrumental variables that are correlated with size but unrelated to recent performance. Finally, McLemore (2019) use fund mergers as shocks to fund size. All these studies report a negative but mostly insignificant relationship. For non-U.S. funds Ferreira, Keswani, Miguel and Ramos (2013) find increasing returns to scale.

In this paper, we re-examine the size-performance relation with a novel identification strategy addressing the endogeneity concern for domestic European funds. Suppose we choose to accept that asset managers are forced to increase their level of activity and have limited opportunities to use in- or outflow to do so. In that case, we introduce variation in the amount of assets under active management. In such a quasi-natural experiment setting, we can identify the size-performance. Our findings suggest that we have decreasing returns to scale. We find that the mean performance estimates in the treated groups are negative relative to the different control groups. The result also holds when we use a more direct measure of marginal bets. We find that the primary source of the underperformance comes from the new bets taken by managers under scrutiny.

Finally, our detailed data allow us to learn how managers behave when pushed to increase their active share. We can investigate every bet of every fund manager. A problem with traditional activity measures such as tracking error and active share is that they do not detect whether funds perform true active management by taking new bets, increase existing ones, or whether they engage in signal-jamming to appear truly active (Brown and Davies (2017) or Cremers, Fulkerson and Riley (2020*a*)). Signal jamming adds tracking error to their returns by taking random bets to generate a false sense of active management. This strategy improves closet indexers' chance to pool with genuinely active funds. Signal jamming implies that closet indexing may be more widespread than indicated by measures that rely on comparing portfolio weights to benchmark weights, such as tracking error and active share.

To separate between signal jamming and more concentrated bets, we use a concentration measure based on the Herfindahl measure (used for example, in Brands, Brown and Gallagher (2005), Kacperczyk, Sialm and Zheng (2005) and Kacperczyk, Nieuwerburgh and Veldkamp (2014)). We find that it is hard to identify more concentrated bets at the total portfolio level. When we split the portfolio into several parts, we find that if managers are forced to increase activity, they take on average more concentrated active bets in stocks they are not already betting on, i.e., they take new bets in unfamiliar stocks. This finding is in accordance with Pollet and Wilson (2008), who found that funds diversify in response to new flows, especially if they operate in relatively illiquid markets.

## 3 Problem diagnosis

In this section, we develop an understanding of the problems with closet indexing and form an outline of the drivers of poor outcomes resulting from the underlying market imperfections.

### 3.1 Closet indexing

The problem with closet indexing is that investors do not get the actively managed funds they pay for and are promised by the fund company. When investors buy active mutual funds, they evaluate investor information, including fee structure.[11] The information lays the foundation for the investors' expectation of active management. If this is incorrect, investors will base investment decisions on a false expectation of getting a more active fund management service than they actually receive (Cremers and Quinn, 2016). Closet indexing does not offer the same ex ante risk profile that investors should expect from genuine active management.

It is hard for retail investors to evaluate the service they receive. The effort of the manager is only partially observable. Even if monitoring is possible, investors may have trouble interpreting the uncovered information. To mitigate moral hazard, we need suitable measures of effort.[12] A natural candidate for an activity measure is the actual outcome of active management, the excess return. However, this is not a reliable measure of the degree of active management. A manager can be very active, but if the bets cancel each other out, the excess return is close to zero. Thus, realized return cannot be used to identify closet index funds. We need measures based on manager effort.[13]

The most used measure for identifying closet indexing is active share. This evaluates the degree of active management for funds relative to the benchmark (Cremers and Petajisto, 2009).[14] An alternative is tracking error. Later we will show that active share and tracking error are highly correlated for European domestic funds. If tracking error is low, then active share is also low, and vice versa. Therefore, we focus on active share as our activity measure.

In order to outperform the market, the active portfolio must differ from the market. Activity is a precursor to superior returns. Since closet index funds charge fees similar to truly active funds, while holding portfolios similar to the index, the net performance of closet index funds is on average lower than the net performance of the much cheaper index funds (ESMA, 2020).

### Motives for closet indexing

To crack down on closet indexing, it is important to understand the motives behind it. Following Berk and Green (2004), the quality of marginal investment opportunities declines as active AUM increases. Under no anti-closet-indexing constraint, managers will choose a (subjective) threshold alpha and closet-index assets when opportunities fall below that point. Gross investor returns are maximized when that threshold alpha is zero.[15]

However, a manager might use a different threshold level. If they are risk or effort averse, they will choose a positive threshold alpha. If they are overconfident, risk seeking (e.g., due to convex

[11] The Key Investor Information Document (KIID) includes the most important information for the investor. Commission Regulation 583/2010 provides a harmonized regime on the form and content of the document, ensuring that information in the UCITS markets is consistent and comparable. KIIDs include sections on objectives and investment policy, risk and reward profile, fees, past performance, and practical information.

[12] There may be different views on the manager's obligation. In the DNB Norge case, the fund claimed in court that the obligation was to use resources to search for bets, not to implement them. We will assume that we can measure the obligation.

[13] In the law there is a separation between the duty to achieve a specific result and the duty of best effort, for more about this, see for example UNIDRIOT Principles.

[14] Active share is defined as one half of the sum of the absolute value of portfolio weights differing from the benchmark.

[15] With this framework, we are different from the assumption regarding that capital is competitively allocated in Berk and Green (2004). In our setting, there are several reasons why this condition does not hold. First, the event window is fairly short. Retail investors typically are very slow with changing their positions. They "suffer" from inertia (see e.g. Bilias, Georgarakos and Haliassos (2010) or Agnew, Balduzzi and Sunden (2003)). Second, tax motives favor not moving in or out of different funds. Before and during the event window, the markets went up, and most investors had large gains on their investments. If these investors had sold their holdings and moved into another fund, they would have to pay tax on gained capital.

incentives), or have a desire to signal skill via activity, they will choose a negative one. Cremers and Quinn (2016) suggested four alternative motives for closet indexing. We describe two of them here, one resulting in a positive threshold alpha and one resulting in a negative one.[16]

If the managers want to preserve their current asset base and are afraid that underperformance might lead to a large outflow from the fund, they can set a positive threshold alpha. It is optimal for investors with a higher active share, but the managers choose a closet indexing strategy. All else equal, larger funds generate more revenue than smaller funds. However, there is an asymmetry in the relationship between flows and performance. While small funds seek to create good returns to grow their asset base, the preference of large funds may be to preserve their current assets and avoid losing badly to the benchmark (Sirri and Tufano, 1998). One way to prevent underperformance is to put much of the fund's assets in the benchmark. While this strategy means the fund will also be unlikely to beat the benchmark, average performance may be enough to maintain a large asset pool, and thus its profitability. Therefore, closet indexing may be a valuable strategy for a risk averse manager seeking to maximize assets under management. If the manager uses a positive threshold alpha, the investor will not benefit from the manager's skill, much like when a good soccer player sits on the bench most of the game. In this case, the investor would be better off if the manager increased the degree of activity.

If the managers have run out of new ideas, they may set a threshold equal to or lower than zero. Assuming that the managers have private information about their skills and that this is lower than what the investors believe, the manager will run out of good ideas more quickly than given by the investor information. Size can be a reason why managers run out of ideas. A large fund has a more limited investment opportunity set consisting of only sufficiently large investments to make a difference. In this case, the investor is not better off with a higher active share.

### The rationale for intervening

From a regulatory perspective, an important element is that mutual funds provide fair and transparent information about how they manage their portfolios. Investors buy active funds for the opportunity to beat the index alternative. However, closet indexing is a significant drag on mutual fund investors' returns, and closet index funds underperform the market and leave investors worse off than other investment choices. Thus, closet indexing potentially poses a regulatory problem for the mutual fund industry.

Cremers and Quinn (2016) develop two approaches to the issue of closet indexing. They suggest a disclosure regime that would incorporate more information, such as active share and an approach where closet indexing is potentially liable under existing laws. In the next section, we will show that both of these approaches were used by regulatory authorities in Scandinavia. In Section 5 we will see that the consequences from regulation can give insight into the motives for closet indexing.

## 4 Intervention design

This section describes the intervention design, and the laboratory that we use to answer the questions, namely active domestic European funds.

[16]We do not expand on the following two motives: 1) Closet indexing is chosen due to the high cost of performing true active management, or 2) closet indexing is chosen due to time-varying active management possibilities.

### 4.1 Policy scrutiny in Europe

Even if most of the literature focus on U.S. mutual funds there are some papers investigating European funds, see for example Banegas, Gillen, Timmermann and Wermers (2013), Ferreira et al. (2013), Cremers et al. (2016) and Leippold and Rueegg (2020). We perform a diff-in-diff study where we compare funds exposed to policy scrutiny in Scandinavia with funds from European countries not exposed to scrutiny. Our treated group is Scandinavian closet index funds since these countries were the first with a regulatory focus on closet indexing in Europe (or the world). They are also the countries with the toughest policy intervention. [17] As a control group, we use closet index funds from European countries where we are unable to identify a meaningful level of scrutiny. Even if there is a selection problem between closet index funds and truly active funds, we will, in a robustness analysis, use truly active funds in Scandinavia as control funds. Below we describe the interventions in detail.

#### Denmark

Even if we set an exact date in Table 1, these scrutiny processes lasts for a good while. Therefore we define an event window of two years. At the beginning of 2014, the Financial Supervisory Authority of Denmark released an analysis of closet indexing in their report for 2013. Using 60% active share and a tracking error of 4% as the limit, they found that 56 out of 188 equity mutual funds had not practiced the active management strategy they marketed in their prospects (Financial Supervisory of Denmark (2013)). When the FSA lowered the limits to 50% and 3% for active share and tracking error, respectively, the number of potential closet indexers went down to 22 funds. Based on the report, the FSA contacted the boards of these funds and requested explanations. Apparently, the Danish FSA was satisfied with the answers as they did not go further with the investigation. However, they imposed requirements for the funds to report active share and tracking error.

#### Norway

In the same period, the Norwegian Financial Supervisory Authority did an extensive report on the level of active management on a subsample of Norwegian mutual funds (Norwegian Ministry of Finance (2015)). Based on their findings, they chose to publicly criticize two funds in November 2014. We have already told the story about DNB Norge. As illustrated in Figure 1 in the Introduction, the fund had an exceptionally low active share. The Norwegian FSA decided to impose a corrective order on DNB, who could either bring the management of the fund in line with the characteristics of true active management, as reflected by the management fee and in the fund's prospectus, or adjust the cost to a level in line with the strategy that was applied. The second fund that the FSA criticized was Nordea Avkastning. They got the same alternatives, either change the level of activity or the fee, even if the activity level was higher than for DNB Norge.[18]

#### Sweden

Also, the Swedish Financial Supervisory Authority analyzed Swedish actively managed mutual funds in 2014. They examined whether the key investors' documents of funds marketed and sold in Sweden

[17] Two articles from Financial Times: Financial times article (2016a) or Financial times article (2016b).

[18] The Financial Supervisory Authority of Norway, available in Norwegian at Finantilsynet.

**Table 1.**
**Country selection**

In this table we present how the countries in Europe are divided into treated, control, or omitted from the sample, respectively. As a starting point, we use the countries included in Ferreira et al. (2013) and Cremers et al. (2016).

| Country | Date and documentation |
|---|---|
| | **Panel I: Treated** |
| Denmark | *September 2014*<br>Danish FSA publish a report showing that 56 out of 188 funds studied were potential closet indexers. More details in text. |
| Norway | *March 2015*<br>Norwegian FSA sends a corrective order to DNB asset management regarding active management. More details in text. |
| Sweden | *October 2015*<br>Finansinspektionen (FI) publish a report on consumer protection in financial markets debating stricter rules.<br>More details in text. |
| | **Panel II: Control** |
| Austria | No scrutiny identified |
| Belgium | No scrutiny identified |
| Finland | No scrutiny identified |
| Poland | No scrutiny identified |
| Portugal | No scrutiny identified |
| Switzerland | No scrutiny identified |
| | **Panel III: Omitted** |
| Italy | *March 2016*<br>The Italian regulator took action against some of the largest investment companies in its home market for mis-selling actively managed funds that closely hugged an index. |
| Netherlands | *May 2016*<br>AFM publish a report on index hugging identifying 7 out of 85 funds investigated to be closet trackers. |
| Germany | *September 2016*<br>BaFin completes its investigation into closet indexing, identifying deficiencies in transparency. |
| France | *March 2017*<br>AMF reminds asset management firms on importance of clarity in the investment objective. |
| United Kingdom | *June 2017*<br>FCA publishes their final report on the Asset Management Market study finding £109 bn invested in closet funds. |
| Luxembourg | *August 2017*<br>CSSF issues a reminder on improving clarity in the "objectives and investment policy" section of the KIID. |
| Spain | *October 2018*<br>CNMV analyzed the existence of these products without reaching any conclusion on how to act.<br>More details. |
| Ireland | *July 2019*<br>Central Bank of Ireland published largest data driven study of industry about closet indexing to date. |

provided accurate and clear information regarding the investment objective and policy. The investigation is presented in the Swedish FSA's annual report on Consumer Protection 2015 (Financial Supervisory of Sweden, 2015). The intervention started a debate on the legal issues related to closet indexing. In 2014 the Swedish Shareholders' Association (Sveriges Aktiesparares Riksforbund) initiated a class action against two mutual funds from one of the largest Swedish banks but decided not to go through with the lawsuit in July 2015.[19] For the Scandinavian countries, Kjørven (2019) analyze how the European legal framework has been applied, and discuss the need for legal measures to ensure that the investors get what they pay for, as protection against closet indexing. In the U.S. mutual fund market Cremers and Quinn (2016) points out the legal issues related to the fact that a large share of

[19]For details about the funds, see case number 2014-11304 (Swedish Ministry of Finance, 2016).

the U.S. mutual fund assets are invested in closet index funds.

### Rest of Europe

Closely following Scandinavia, the European Securities and Market Authorities (ESMA (2016)) wrote an extensive report based on more than 2,600 European funds to map the presence of closet indexers. Their findings suggest that between 5% and 15% of the sample funds were potential closet indexers. As we see from Table 1, several other countries performed their own investigations. To our knowledge, none of these have resulted in any legal claims against funds.

### Event design

To conduct an event study, we need to follow the mutual funds before, during, and after the interference. We let the window of policy scrutiny start in January 2014 and last until December 2015. The pre-event period is the two years before January 2014, and the post-event period is the four years after, starting from January 2016 until December 2019.

We have two groups, treated and control funds. For the treated funds, we cannot always directly attribute the consequences of scrutiny to the behavior of the asset managers. In that sense, the DNB Norge case is an exception. However, based on the policymakers and financial authorities' interventions and the following interest in media, we assume that the funds that were in danger of being labeled as a closet index fund were rethinking their active management strategy. Furthermore, we assume the intervention's effect occurred during the two-year window from January 2014 to January 2016.

To identify a control group that is as free as possible from scrutiny, we have sorted all the countries in Europe except Scandinavia into two groups, one without scrutiny and one where we have identified it. In Panel II: Control, we list the countries that constitute our control group, and in Panel III: Omitted, we list the countries that have exhibited scrutiny towards the fund industry. We also link to documentation that indicates the policy scrutiny.

## 4.2 Data and summary statistics

This section describes the data, how we construct our treatment and control groups, and presents summary statistics.

### Sample selection

The dataset is built using two primary databases: Morningstar Direct and Lipper Fund Database. The data ranges from January 2010 through December 2019. Our focus is on domestic long-only equity funds.[20] The fund data include monthly assets under management and gross and net returns.[21] The expense ratio is the monthly price of active management and is calculated using the net and gross returns. We download monthly holdings for each fund to calculate active share.[22] Details on the sample selection and raw data are presented in Appendix B.1.

[20] We define domestic funds as funds with an investment area equal to the home country of the investment company and a domestic primary prospectus benchmark.

[21] Returns are in the local currency, while assets under management are in USD to have a common currency for comparison across countries.

[22] We have filled in with Morningstar for funds with missing data in Lipper.

For each fund, we define a benchmark for the calculation of active share and performance evaluation. Benchmark constituents, weights, and returns come from Datastream. We use the Lipper technical benchmarks whenever these are available. This benchmark assignment method minimizes the concern that funds strategically choose benchmarks that might not accurately reflect their actual investment style. For active funds, in general, categorization is complicated but is minimized by investigating domestic funds. We focus on domestic funds since these give us the most homogeneous group of funds both across and within countries.[23] Details on the benchmark selection and data are provided in Appendix B.2.

Both survivors and non-survivors are included to avoid survivorship bias. For a sufficient number of observations to get meaningful inference in the regressions, we exclude funds with less than 12 months of data in either the pre- or post-event period. After the exclusions, we are left with a sample consisting of 353 funds, from which we define treatment and control groups based on active share. The full sample is presented in Table B.2 in Appendix B.3.

To evaluate factor returns, we collect size and style portfolio returns from MSCI. Calculations of the return factors are detailed in Appendix B.4. We present a list of all variables, divided into outcome and controls in Table B.3 in Appendix B.5.

#### Defining closet indexing

There are several alternatives for how to define a closet index fund. The two most common measures are active share and tracking error. We choose to focus on active share. In Appendix B.6 we show that, for domestic funds, the factor structure is weak and so the correlation between active share and tracking error is high. To define a fund as a closet index fund, we set two alternative cutoff points, 40% and 50%, with funds being classified at the start of the event window. Furthermore, we assume that regulatory or legal risk increases with lower levels of activity.

### 4.3 Summary statistics

In Table 2 we report summary statistics for the domestic European closet index funds. We report active share and other fund characteristics before the event window for the treated sample (Scandinavian funds) and the control sample (European, non-Scandinavian funds). We denote the benchmark-adjusted performance as alpha. We agree with Berk and van Binsbergen (2015) that a tradable index-based adjustment is likely to adjust for fund style and risk more precisely than the loadings on risk factors. More details on the benchmarks are provided in Appendix B.2. In robustness tests factor-adjusted alpha are also analyzed, with details on factor calculations provided in Appendix D.5. The expense ratio is the price, presented as a percentage, that the investors pay for professional money management. Net alpha is the gross alpha minus the expense ratio.

In Column 1-3, we report the numbers for the cutoff at 40%. Out of 79 funds, 46 are treated, and 33 are controls. The average active share is about 30%. Monthly average gross alpha for the two years before the event is in the range of five to nine basis points, where the Scandinavian funds perform best. After fees, the average alpha is slightly negative. These findings are similar to other results in

[23]Their mandates are similar. They have a comparable size in investment universes (defined by the benchmark index), and the funds make up approximately the same size of the local stock exchange.

**Table 2.**
**Summary statistics: Closet index funds**

This table presents summary statistics for the sample of closet index funds in Scandinavia (Denmark, Norway, and Sweden) and the rest of Europe. Active share, assets under management, and fund age is as of December 31, 2013 before the event window. Expense ratio, gross and net alpha estimates are means over a two-year window before the event.

| Limit active share | 40% | | | 50% | | |
|---|---|---|---|---|---|---|
| Group | Treated | Control | Diff | Treated | Control | Diff |
| Number of funds | 46 | 33 | | 75 | 47 | |
| Active share (%) | 30.40 | 31.40 | −1.00 | 35.90 | 35.20 | 0.67 |
| Gross alpha (%) | 0.09 | 0.05 | 0.04 | 0.10 | 0.09 | 0.01 |
| Expense ratio (%) | 0.10 | 0.10 | 0.00 | 0.10 | 0.10 | 0.00 |
| Net alpha (%) | -0.01 | -0.05 | 0.04 | 0.00 | -0.01 | 0.01 |
| AUM (million USD) | 589 | 245 | $344^{***}$ | 463 | 252 | $210^{**}$ |
| Fund age (years) | 17.1 | 16.9 | 0.21 | 16.00 | 16.30 | −0.30 |

the literature, with a small positive alpha before fees and zero after. The treated and control sample is similar along all dimensions except for assets under management. We find that the treated funds are larger than the control funds. We will use econometric techniques, covariates and fixed-effects, to correct for these differences. Another alternative is to use a matching procedure. Based on findings in the mutual fund literature, variables such as skill (alpha), size (AUM), and fund age are reasonable candidates for matching, which we return to in Section 5. It is worth noting that, compared, for example, to a U.S. fund, these funds are small in terms of their management teams and organization scope but large relative to their investment space.

In Column 4-6, we report the numbers for the 50% limit. In this case, there are more closet index funds. From a pool of 122 funds, 75 are treated, and 47 are controls. The average active share is close to 35 percent. Monthly average gross alpha for the two years before the event is around ten basis points, slightly higher than for the 40% cutoff. Again the Scandinavian funds perform best. After fees, the average alpha is marginally negative. Again, the treated and control sample are similar along all dimensions except for size.

# 5 Impact assessment

We study the intervention's effect in two "stages". First on the level of activity and fees, and next its impact on value creation. For active share and fees, our primary interest lies in what happens while the funds are under scrutiny, but for value creation it is the consequences of the potential new behavior that is interesting. We will show that the second stage's impact works through changes in activity level, i.e., funds under scrutiny change activity level, which again affects value creation. The later stage is a direct test of how a change in activity influences value creation, i.e., decreasing returns to scale.

## 5.1 First stage: Impact on active share and fees

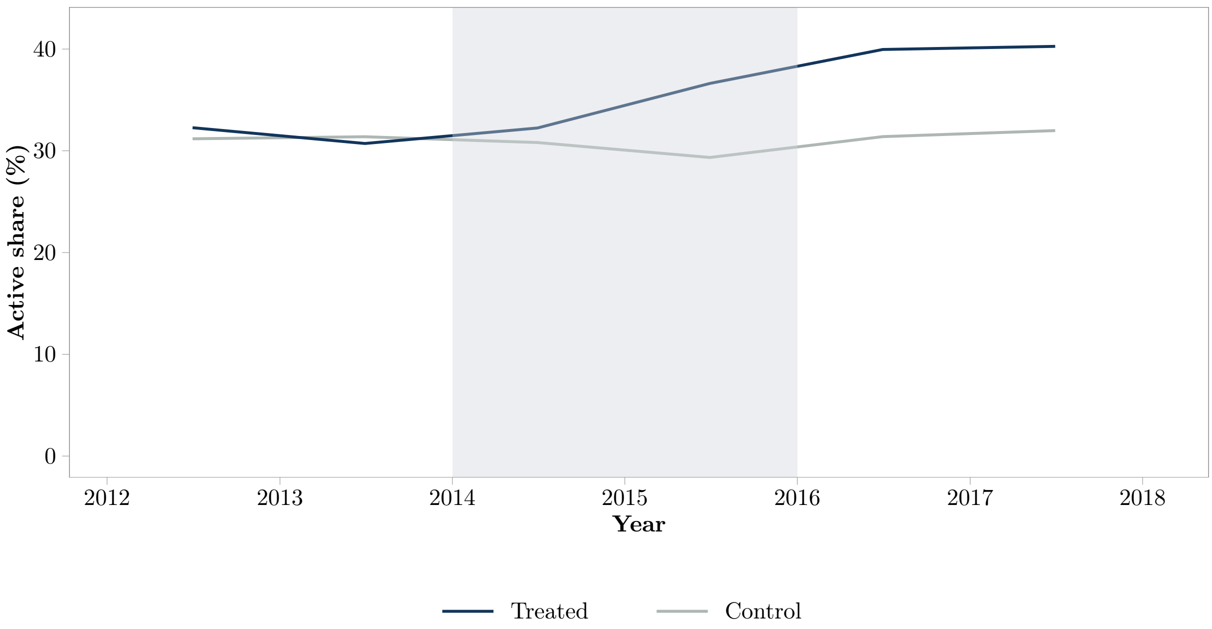
We have already hypothesized that the funds will choose the latter in the choice between reducing fees and increasing activity. In this section, we empirically test what happens under scrutiny for both alternatives and start with active share.

### Active share

We do a visual analysis before we perform more formal tests. Figure 2 illustrates the development of mean active share for treated and non-treated closet index funds over time, and clearly show that funds under scrutiny increase active share more than unaffected funds. Before the event, the funds' activity is similar, but the active share for treated funds increases more than for non-treated funds during the event window. An essential prerequisite for a diff-in-diff analysis is a pre-trend evaluation. Both in the visual illustration and a formal test in Appendix C.1, we document that the trends are parallel.

**Figure 2.**
**Development of active share**

This figure presents the progression in active share for closet index funds from January 2012 to December 2017. The time series are the annual averages of monthly cross-sectional group-wise means, and the shaded area highlights the event window.




Formal tests for the impact of scrutiny on active share is presented in Table 3. The detailed analyses show that scrutiny leads to a higher level of activity, with the first regression showing the actual effect, i.e., how large the average difference is between the groups. To minimize endogeneity concerns, we formally test the effect of scrutiny in the last columns for each group using fixed effects (fund and time) panel regressions. Results are reported using a two-years pre- and post-event window.

In Column 1-5, we present the results for the 40% cutoff point. In Column 1, we calculate the average difference between a treated and non-treated fund for the two years after the event. Using "collapsed" data, i.e., averages over time, we avoid any potential autocorrelation between the monthly observations (Bertrand, Duflo and Mullainathan, 2004). As Table 2 shows, the difference in the pre-period is small enough to interpret the difference in the post-period as a treatment effect. Moreover, to obtain actual differences, controls are not included in this OLS regression. As such the results show that, on average, treated funds have an 8 percentage points higher active share than non-treated funds after the event. Relative to the active share before the event, the increase is around 26%(= 7.9%/30.4%). In Column 2, we use a matched control group based on skill, fund size and fund age. For more details about the matching, see Appendix C.2. The estimate using the matched sample is close to our first analysis. Column 3 reports the result from a diff-in-diff analysis where we compare the difference from before to after between treated and control funds. With small pre-event differences, it is not surprising that the diff-in-diff estimate is equalling our previous estimates, around 8 percentage points.

**Table 3.**
**Effect from intervention on active share**

This table present tests on the effect the intervention had on active share for closet index funds. We use 40% and 50% as cutoff points to classify funds as closet indexers. The model in Column 1-2 and 6-7, labeled Diff post, is a cross-sectional OLS model, where we compare the collapsed group-wise means before and after the event. The difference between the estimates are due to the underlying sample, where we in Column 1 and 6 use the actual sample, whilst in Column 2 and 7 we use a matched sample based on one-year pre-event average gross alpha, assets under management, and fund age. DiD is a difference-in-differences model and FE are panel data regressions with both time and fund fixed effects. Standard errors are presented in parentheses. In Column 4 and 9 standard errors are clustered at the fund level, whilst in Column 5 and 10 they are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: $^{***}$p<0.01, $^{**}$p<0.05, $^{*}$p<0.1.

| | 40% | | | | | 50% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diff post | Diff post | DiD | FE | FE | Diff post | Diff post | DiD | FE | FE |
| Scrutiny | 8.211$^{***}$ | 7.850$^{***}$ | | | | 6.206$^{***}$ | 5.873$^{***}$ | | | |
| | (2.068) | (2.063) | | | | (2.074) | (2.085) | | | |
| Scrutiny · Post | | | 7.873$^{***}$ | 7.873$^{***}$ | 8.892$^{***}$ | | | 5.057$^{*}$ | 5.780$^{***}$ | 7.087$^{***}$ |
| | | | (2.702) | (2.172) | (2.697) | | | (2.723) | (1.835) | (2.115) |
| Controls | No | No | No | No | Yes | No | No | No | No | Yes |
| Fund FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Time FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Sample | Actual | Match | Actual | Actual | Actual | Actual | Match | Actual | Actual | Actual |
| N | 79 | 77 | 158 | 3,790 | 3,693 | 122 | 119 | 244 | 5,704 | 5,513 |
| Adjusted $R^2$ | 0.159 | 0.151 | 0.165 | 0.675 | 0.716 | 0.062 | 0.056 | 0.078 | 0.714 | 0.735 |

In Column 4-5, we perform panel data regressions with both time and fund fixed effects. The fund fixed effect soaks up any variation in active share due to cross-sectional differences in fund characteristics. The identification comes from variation over time within a fund, not from variation across funds. In Column 5 we add control variables, described in Appendix B.5. The interaction effect between after and funds under scrutiny is similar to previous analyses, around 8 percentage points. The estimate is large and significant. In the continuation of the paper, we report one regression where we cluster standard errors on the fund level and one where we cluster independently along the fund and time dimension. For most of the outcome variables, such as active share or expense ratio, allowing a time dimension does not significantly affect inference. But it will be more relevant for active returns since, even after removing a time fixed effects, mutual funds may have returns that are not independent of one another within time periods. Therefore, we will elaborate on two-way clustering when we come to the value creation part in Section 5.2.

In Column 6-10, we repeat the same specifications as we did in Column 1-5, but for the 50% cutoff point. The smaller effects in Column 6-10 are what we would expect to see given that the degree of scrutiny is negatively related to the size of active share. For example, the diff-in-diff estimate is about 5 percentage points or around 14%(= 5.1%/35.9%) higher than the active share before the event. Thus, independent of the cutoff level, the funds under scrutiny have increased the activity level relative to non-scrutinized funds.

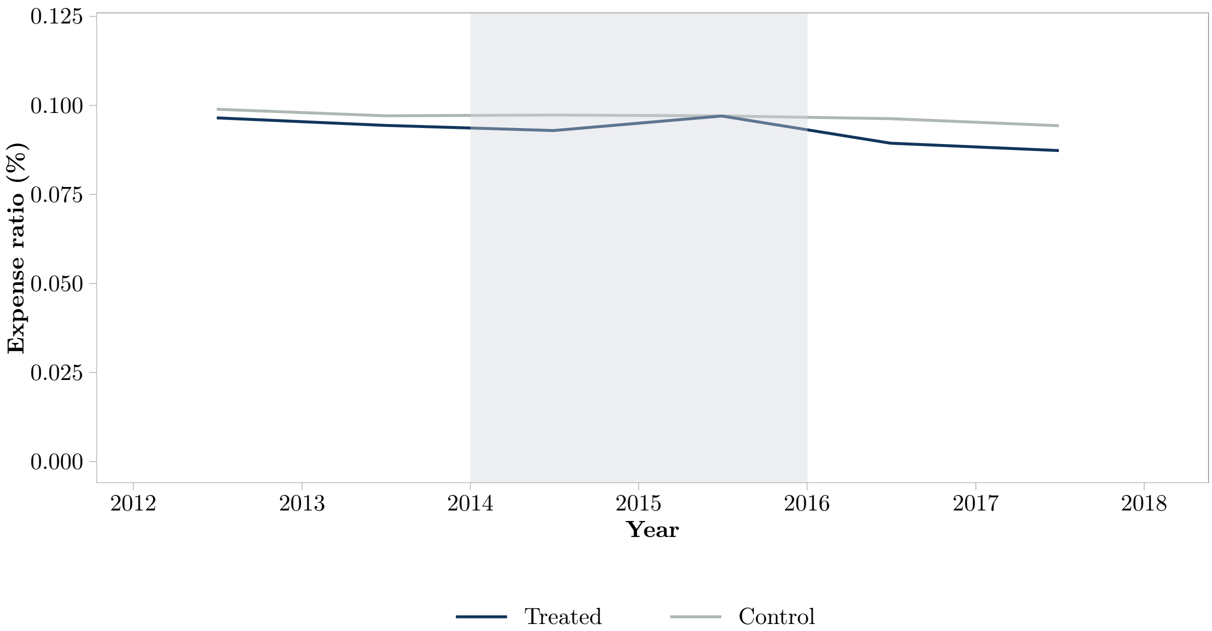
A potential concern is that the change in activity does not come from real bets but from managers only adding noise to their returns, known as signal jamming (Brown and Davies (2017) and Cremers et al. (2020*a*)). A problem with the activity measures used so far is that they do not detect whether funds engage in signal jamming to appear truly-active. A measure that can identify whether the funds are randomly selecting stocks or taking more concentrated bets is the Stock concentration index ($SCI$). This is a stock-level version of the industry concentration index developed by Kacperczyk et al. (2005). In Appendix C.3, we form analyses with $SCI$ as the dependent variable. Regarding the overall active portfolio, the evidence is mixed on whether forced managers perform truly active management or signal jamming, but in Section 6.1 Best ideas, where we split their portfolios into different parts, it becomes clear that mangers under scrutiny take larger bets in new stocks than non-scrutinized managers.

### Fees

We have hypothesized that fund managers under scrutiny choose to not change the fee level. Again, a visual analysis precedes formal testing. In Figure 3 the development of the mean expense ratio for treated and non-treated funds is illustrated. For both groups there is a decreasing trend, which is in line with the overall trend for active mutual funds (Morningstar, 2019). Intra-group discrepancies are small and stable, meaning that funds put under scrutiny have not lowered their costs more than unaffected funds. For a formal pre-trend analysis, see Appendix C.1.

**Figure 3.**
**Development of fund fees**

This figure presents the progression in expense ratio for closet index funds from January 2012 to December 2017. The time series are the annual annual averages of monthly cross-sectional group-wise means, and the shaded area highlights the event window.




In Table 4 results from formal tests regarding the impact of scrutiny on fees are presented. The detailed analyses show that scrutiny did not lead to a relatively lower fee level. As in Table 3 we report actual effects first, before formally estimating the scrutiny effect. In similar manner as before, results for the 40% cutoff point is presented in Column 1-5 and in 6-10 for the 50% level. Every estimate is small and insignificant, thereby confirming our hypothesis that managers under scrutiny choose to increase activity over updating investor information and fees.

**Table 4.**
**Effect from intervention on expense ratio**

This table present tests on the effect the intervention had on expense ratio for closet index funds. We use 40% and 50% as cutoff points to classify funds as closet indexers. The model in Column 1-2 and 6-7, labeled Diff post, is a cross-sectional OLS model, where we compare the collapsed group-wise means before and after the event. The difference between the estimates are due to the underlying sample, where we in Column 1 and 6 use the actual sample, whilst in Column 2 and 7 we use a matched sample based on one-year pre-event average gross alpha, assets under management, and fund age. DiD is a difference-in-differences model and FE are panel data regressions with both time and fund fixed effects. Standard errors are presented in parentheses. In Column 4 and 9 standard errors are clustered at the fund level, whilst in Column 5 and 10 they are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

| | 40% | | | | | 50% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diff post | Diff post | DiD | FE | FE | Diff post | Diff post | DiD | FE | FE |
| Scrutiny | −0.008 | −0.007 | | | | 0.005 | 0.006 | | | |
| | (0.010) | (0.010) | | | | (0.008) | (0.008) | | | |
| Scrutiny · Post | | | −0.005 | −0.004 | 0.006 | | | 0.001 | 0.002 | 0.004 |
| | | | (0.014) | (0.004) | (0.007) | | | (0.011) | (0.003) | (0.006) |
| Controls | No | No | No | No | Yes | No | No | No | No | Yes |
| Fund FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Time FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Sample | Actual | Match | Actual | Actual | Actual | Actual | Match | Actual | Actual | Actual |
| N | 79 | 77 | 157 | 3,812 | 3,744 | 122 | 119 | 243 | 5,887 | 5,730 |
| Adjusted $R^2$ | 0.007 | 0.006 | 0.007 | 0.907 | 0.909 | 0.004 | 0.005 | 0.005 | 0.904 | 0.905 |

To sum up, policy scrutiny on closet index funds has impacted active management. In the choice between reducing fees or increase activity, they have done the latter. This supports the notion that the managers "forced" themselves to increase activity. In the next section, we investigate the consequences of increased activity.

### 5.2 Second stage: Impact on value creation

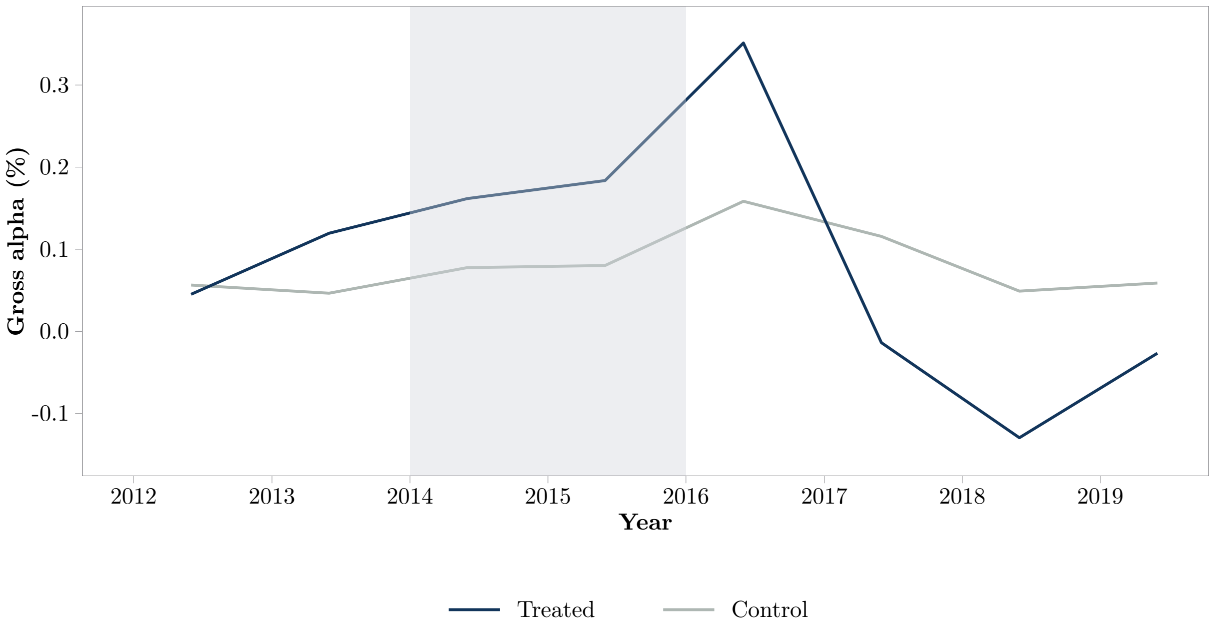
In this section, we investigate the impact on value creation, by first showing the actual effects, i.e., how much value the group of scrutinized funds have created relative to the unaffected group of funds, before formally testing if the effect on value creation is statistically significant. Having demonstrated small and similar effects on fees between treated and control funds, there are grounds to suspect a negligible difference in results for gross and net alpha. Consequently, we focus on the former, but confirm that using the latter does not alter any conclusions.

There are two ways to measure value creation, either by alpha or value added. Following the arguments in Berk and van Binsbergen (2015) and Berk and van Binsbergen (2016) value added for a period (before fees) can be viewed as the total value the funds extract from the capital markets. Value added is the product of gross active return, alpha, multiplied with the fund's AUM at the end of the previous period. Gross value added can be divided into two: the amount of money the manager takes home for himself (fees) and the amount given to investors. The latter is named net value added. We will focus on alpha in the main text but also show results for value added. Again, our conclusions remains largely the same, with any interesting differences highlighted in the main text.

Again, we start with a visual analysis before presenting formal tests. Figure 4 illustrates the development of mean gross alpha for treated and non-treated closet index funds. In the pre-event period, it seems that treated funds perform somewhat better than non-treated funds. An important prerequisite for a diff-in-diff analysis is a pre-trend evaluation, which is documented in Appendix D.1. To evaluate the impact of the intervention, we use a four-year period and show that funds under scrutiny underperform non-scrutinized funds.

**Figure 4.**
**Development of gross alpha**

This figure presents the progression in gross alpha for closet index funds from January 2012 to December 2019. The time series are the annual annual averages of monthly cross-sectional group-wise means, and the shaded area highlights the event window.




In Table 5, we present the formal tests for the impact scrutiny has on gross alpha. The first regressions show the actual effect on returns, i.e., how much the treated group, on average, has performed against the control group. The last regressions test if the effect of scrutiny is statistically significant. To minimize the endogeneity problem, we include both fund and time fixed effects.

In Column 1-5, we present results for the 40% active share cutoff. In Column 1, the average difference between treated and non-treated funds for the four-year period after the event is reported. This shows that, on average, scrutinized funds that increased their active share have a 6 basis points lower active return per month than unaffected closet index funds. In annual terms, this difference becomes 70 basis points or almost 3% over four years. Had the regulatory authorities followed the alternative with updating investor information and reducing fees, investors could have fared even better. Assuming a fee reduction of half of what DNB had to pay back, the annual increase in value creation would have been $1.1\%(= 0.7\% + 0.4\%)$ or 4.4% higher over the period.

In Column 2, we show the estimate for the difference when we match the treated closet index funds with control funds based on skill, size and fund age. The estimate is 5 basis points per month or 60 basis points annually. In Column 3, we compare the difference between before and after for treated and control funds. The diff-in-diff is about 9 basis points per month or 1.1% annually. Again the estimate is significant. One way of interpreting this result is that an investor that was long the average control fund and short the average treated fund before the intervention and long the treated fund and short the control fund after intervention would have lost 1.1% annually.

**Table 5.**
**Effect from intervention on gross alpha**

This table present tests on the effect the intervention had on gross alpha for closet index funds. We use 40% and 50% as cutoff points to classify funds as closet indexers. The model in Column 1-2 and 6-7, labeled Diff post, is a cross-sectional OLS model, where we compare the collapsed group-wise means before and after the event. The difference between the estimates are due to the underlying sample, where we in Column 1 and 6 use the actual sample, whilst in Column 2 and 7 we use a matched sample based on one-year pre-event average gross alpha, fund age and assets under management. DiD is a difference-in-differences model and FE are panel data regressions with both time and fund fixed effects. Standard errors are presented in parentheses. In Column 4 and 9 standard errors are clustered at the fund level, whilst in Column 5 and 10 they are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: $^{***}$p<0.01, $^{**}$p<0.05, $^{*}$p<0.1.

| | 40% | | | | | 50% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diff post | Diff post | DiD | FE | FE | Diff post | Diff post | DiD | FE | FE |
| Scrutiny | $-0.059^{**}$ | $-0.050^{*}$ | | | | $-0.074^{**}$ | $-0.072^{**}$ | | | |
| | (0.027) | (0.026) | | | | (0.029) | (0.028) | | | |
| Scrutiny · Post | | | $-0.093^{**}$ | $-0.079^{*}$ | $-0.529^{**}$ | | | $-0.090^{**}$ | $-0.074^{**}$ | $-0.396$ |
| | | | (0.045) | (0.041) | (0.265) | | | (0.045) | (0.037) | (0.241) |
| Controls | No | No | No | No | Yes | No | No | No | No | Yes |
| Fund FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Time FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Sample | Actual | Match | Actual | Actual | Actual | Actual | Match | Actual | Actual | Actual |
| N | 79 | 77 | 157 | 5,492 | 5,427 | 122 | 119 | 243 | 8,410 | 8,258 |
| Adjusted $R^2$ | 0.047 | 0.034 | 0.010 | 0.096 | 0.106 | 0.045 | 0.044 | 0.011 | 0.128 | 0.131 |

In Columns 4 and 5, we perform fund and time fixed effects panel regressions. The fund fixed effects soak up any variation in gross alpha due to cross-sectional differences in fund characteristics and time fixed effects remove common variation due to time. In Column 4, we exclude control variables. Then the interaction effect between after and funds under scrutiny is similar to previous analyses, with a significantly negative estimate of around 9 basis points per month or 1.1% annually. In Column 5, control variables are added yielding a much higher estimate of 0.5% per month. The variables are described in Appendix B.5. A discussion on clustered standard errors follows below.

In Column 6-10, the cutoff point is 50%. Given the lower increase in active share for this cutoff, it is natural to expect less negative results than for the sample of closet index funds at the 40% active share level. What we find is in stark opposition to this, with point estimates very similar to those for the 40% level. To reiterate, funds under scrutiny have a lower alpha relative to non-scrutinized funds.

Below we perform five robustness tests related to the gross alpha estimate, and acknowledge that there are just fundamental limits to how much one can learn from comparing the returns of two groups of funds over four years. However, the best we can do is use the most well-known statistical methods to draw useful inferences about returns.

### Clustered standard errors

First, we discuss potential clustered standard errors. Accurate standard errors are a fundamental component of statistical inference, but this issue seems to be in development, and it is not entirely clear what the best solutions are regarding fund returns.[24] In this paper, we report one regression with fund-level clustered standard errors and one with standard errors independently clustered on the fund and time level. As mentioned in Section 5.1, allowing for a time dimension in outcome variables such as active share and expense ratio does not affect inference significantly. For active returns on the other hand it will be relevant, because even after removing time fixed effects, mutual fund returns might still exhibit an intra-time dependence structure. Hence, we elaborate on two-way clustering in Appendix D.2, where Table D.2 shows how that the standard errors increase when we go from one-way fund clustering to two-way fund and time clustering.

### Placebo tests

Secondly, we perform placebo tests. A potential concern is that the estimated impact of regulatory scrutiny is either merely a random effect or captures some spurious correlation(s) with omitted variables. If this were the case, we should obtain the same results independent of the assignment of treatment and control observations. We test this possibility through a placebo test where we randomly assign funds to treatment and control groups, keeping the ratio of treated to non-treated funds identical to the original sample (see Table 2). Using these randomly assigned groups, we estimate the Diff-in-diff model presented in Columns 3 and 8 in Table 5. We repeat this exercise for 1000 estimations and report the resulting $\beta$ coefficients on the Scrutiny $\cdot$ Post interaction term in a histogram in Figure D.1 in Appendix D.3. We find a significantly negative effect (5% confidence level) only for 1 of the 1,000 trials (0.1%) when using 40% as the limit on active share. For the sample using 50% as a limit, the corresponding number of trials with a significantly negative estimate is 13 (1.3%). Thus, only 0.1% (40% limit) and 1.3% (50% limit) of the estimated $\beta$ coefficients on the Scrutiny $\cdot$ Post interaction term are equal to or smaller than the coefficient estimated using the original sample (-0.093 and -0.090, represented by solid vertical lines in the figure). These results reassure us that our tests capture the treatment effect of regulatory scrutiny on fund value creation and not some random effect or omitted variable.

### Net alpha

Third, we investigate net alpha instead of gross alpha. So far, we have assumed that the effect on net alpha is the same as for gross alpha. This assumption is based on documented small fee effects. In Table D.3 in Appendix D.4 we verify our assumption that the results holds also when net alpha is the dependent variable.

### Factor adjusted alpha

Fourth, factor adjusted alphas are analyzed. Limiting our scope to comparing European domestic funds, we do not expect any of the funds to have a particular factor style. The high correlation between

[24] For example, the recent well-published papers Pástor et al. (2015), Pastor and Vorsatz (2020), and Cremers, Fulkerson and Riley (2020*b*), use various ways of clustering.

active share and tracking error indicates a weak factor structure. Often the funds with a factor style have a broader investment universe than a single European country. Consequently, our focus has been on adjusting for risk given by the chosen benchmark. It is also this excess return the investor has earned. However, as a robustness test, we adjust alpha for domestic factor risk and perform the same regression analyses as in the other tables.[25] In Appendix D.5 one can see that the estimates have the same sign but slightly lower magnitude after adjusting for factor exposure. Thus, the negative active returns are caused by a combination of poor stock picking and factor exposure.

**Value added**

Finally, we perform the same analyses as above, but now with value added as the dependent variable. Remember that value added can be interpreted as a value weighting of alpha and can be viewed as the total value the funds extract from the capital markets. In Table D.5 in Appendix D.6 we perform the same analyses as in Table 5 above, but now with value added as the dependent variable. Again, we confirm our previous findings. Even if the statistical significance is slightly lower, funds under scrutiny extract less value from the stock market than funds not under scrutiny. More precisely, treated funds with a cutoff point at 40% have destroyed, on average, 0.10 million USD compared to the control funds per month for the four years after the event. Recall that these are actual numbers. The value treated funds have destroyed, is larger with a cutoff point at 40% than with a cutoff at 50%.

Our analyses show that scrutinized closet index funds underperform non-scrutinized closet index funds. Relating these findings to the motives for closet indexing supports the motive for the manager running out of ideas for new successful bets. Regulation can either directly or indirectly force managers to increase activity. This makes managers invest in stocks in a sub-optimal manner leading them to destroy investor value. In light of this, the best alternative for investors, conditional on supervisory intervention, is an update in investor information according to the realized active strategy. This should be followed by a reduction in fees.

# 6 Robustness tests

Up till now, we have shown that closet index funds under scrutiny increased activity and have lower value creation than funds not under scrutiny. Next, in Section 6.1, we dig deeper into what leads to deficient value creation. In Section 6.2, we discuss whether we can say something about decreasing active returns to scale based on our analyses. Finally, in Section 6.3, we examine how robust our results are for an alternative control group, namely truly active Scandinavian funds.

## 6.1 Digging deeper: Best ideas

So far, we have compared two active portfolios where one is exposed to scrutiny, and one is not. However, it might be that the effect of scrutiny only influences parts of the active portfolio. A better measure will identify the marginal portfolio changes. To achieve this, we split the active portfolios into subgroups based on the bets' size. We develop three sub-portfolios, one with the largest bets in absolute value, one with medium bets, and one with the smallest bets. We refer to the largest bets

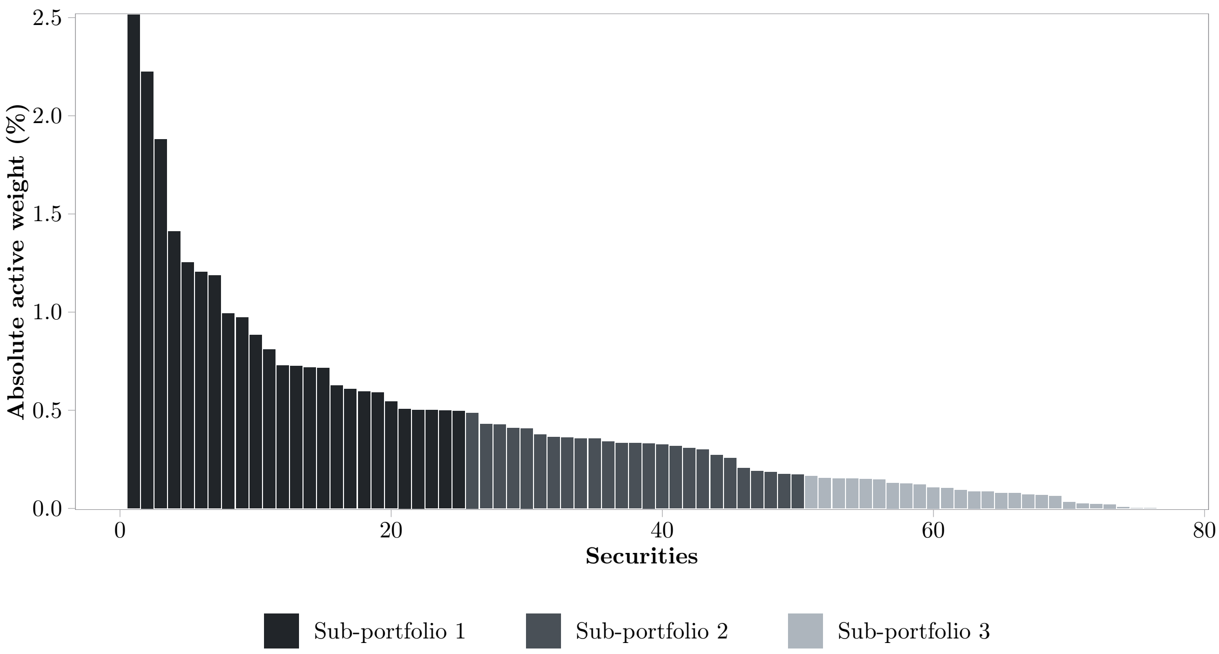
[25]For details, see Appendix B.4.

as the "best ideas", inspired by Pomorski (2009) and Miguel Anton and Polk (2020). Thus, all the existing bets are mapped into one of the three sub-portfolios based on the active weight's absolute size just before the event window.

To illustrate this categorization, we show the distribution of all active weights for DNB Norge at the end of December 2013 in Figure 5. Lastly, we also add a fourth sub-portfolio. This consists of all new bets the fund performs during the period. These bets can come from a change in weight from neutral to positive or negative relative to the benchmark weight or from a new bet in a stock not part of the benchmark. Thus, we can identify whether the managers change the initial active weights or whether they take on new bets in stocks that they do not have a position in.

**Figure 5.**
**Active weights for DNB Norge in December 2013**

This figure present the active weights for DNB Norge in December 2013. Stocks with the largest absolute active weights are in portfolio 1 (best ideas) and those with the smallest absolute active weights are in portfolio 3.




In Table 6 we present our findings in three panels. We start with active share as the dependent variable in Panel I. We present the diff-in-diff estimate for the four portfolios for both the cutoff at 40% and 50%. The estimate is equivalent to Column 3 and 8 in Table 3, but with a different dependent variable. We find that it is in the new, not old stocks, that the scrutinized funds increases their active share.[26] This applies to both the group of funds with an active share below 40% and 50%.

In Panel II, SCI is the dependent variable. We present the diff-in-diff estimate for the four portfolios for both the cutoff at 40% and 50%. The estimate is equivalent to Column 3 and 8 in Table C.3, but with a different dependent variable. We find that scrutiny makes funds take new and more concentrated bets. Finally, in Panel III, gross alpha is the dependent variable. We present the diff-in-diff estimate for the four portfolios for both the cutoff at 40% and 50%. The estimate is equivalent to Column 3 and

[26] Note that the sum of the characteristics we measure for the sub-portfolios is not exactly the same as for the total portfolio, as shown in Section 5. This is because our definition of the sub-portfolios is limited to capture the changes in stocks listed in December 2013 and does not capture all newly listed companies after the event window.

**Table 6.**
**Effect from intervention on sub-portfolios**

This table present tests on the effect the intervention had on active share (panel I), security concentration index (SCI, panel II) and gross alpha (panel III) in closet indexers' sub-portfolios. We use 40% and 50% as cutoff points to classify funds as closet indexers. All results are from a difference-in-differences model equivalent to the one presented in Column 3 and 8 in table 5 in section 5.2. Standard errors are presented in parentheses. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

**Panel I: Active share**

| | 40% | | | | 50% | | | |
|---|---|---|---|---|---|---|---|---|
| | New stocks | Sub-portfolio 1 | Sub-portfolio 2 | Sub-portfolio 3 | New stocks | Sub-portfolio 1 | Sub-portfolio 2 | Sub-portfolio 3 |
| Scrutiny · Post | 3.384*** | 0.586 | 0.303 | 0.726 | 3.242*** | −0.758 | −0.049 | 0.843 |
| | (0.520) | (2.731) | (0.832) | (0.585) | (0.500) | (2.603) | (0.840) | (0.569) |
| Controls | No | No | No | No | No | No | No | No |
| Fund FE | No | No | No | No | No | No | No | No |
| Time FE | No | No | No | No | No | No | No | No |
| Sample | Actual | Actual | Actual | Actual | Actual | Actual | Actual | Actual |
| N | 155 | 158 | 158 | 158 | 237 | 244 | 244 | 244 |
| Adjusted $R^2$ | 0.606 | 0.053 | 0.136 | 0.279 | 0.521 | 0.027 | 0.079 | 0.146 |

**Panel II: Security concentration index (SCI)**

| | 40% | | | | 50% | | | |
|---|---|---|---|---|---|---|---|---|
| | New stocks | Sub-portfolio 1 | Sub-portfolio 2 | Sub-portfolio 3 | New stocks | Sub-portfolio 1 | Sub-portfolio 2 | Sub-portfolio 3 |
| Scrutiny · Post | 0.099*** | −0.038 | 0.063 | 0.034 | 0.106*** | −0.601 | 0.023 | 0.024 |
| | (0.029) | (0.432) | (0.051) | (0.028) | (0.026) | (0.707) | (0.054) | (0.041) |
| Controls | No | No | No | No | No | No | No | No |
| Fund FE | No | No | No | No | No | No | No | No |
| Time FE | No | No | No | No | No | No | No | No |
| Sample | Actual | Actual | Actual | Actual | Actual | Actual | Actual | Actual |
| N | 155 | 158 | 158 | 158 | 237 | 244 | 244 | 244 |
| Adjusted $R^2$ | 0.286 | 0.001 | 0.060 | 0.144 | 0.304 | 0.008 | 0.038 | 0.017 |

**Panel III: Gross alpha**

| | 40% | | | | 50% | | | |
|---|---|---|---|---|---|---|---|---|
| | New stocks | Sub-portfolio 1 | Sub-portfolio 2 | Sub-portfolio 3 | New stocks | Sub-portfolio 1 | Sub-portfolio 2 | Sub-portfolio 3 |
| Scrutiny · Post | −0.038** | 0.011 | −0.019 | 0.011 | −0.070** | 0.017 | −0.028* | −0.018 |
| | (0.017) | (0.027) | (0.017) | (0.011) | (0.033) | (0.030) | (0.016) | (0.023) |
| Controls | No | No | No | No | No | No | No | No |
| Fund FE | No | No | No | No | No | No | No | No |
| Time FE | No | No | No | No | No | No | No | No |
| Sample | Actual | Actual | Actual | Actual | Actual | Actual | Actual | Actual |
| N | 133 | 158 | 151 | 139 | 207 | 243 | 232 | 211 |
| Adjusted $R^2$ | 0.048 | 0.028 | 0.010 | 0.014 | 0.040 | 0.018 | 0.013 | 0.028 |

8 in Table 5, but with a different dependent variable. We find that the relative negative alpha comes from these new concentrated bets. Thus, the main source of underperformance from closet index funds found in Section 5.2 stems from the new bets.

In sum, relative to active managers not put under scrutiny, treated funds do not create excess return on their best ideas nor on bets that they are already familiar with, but in new stocks. We can only speculate, but a potential explanation is that closet index managers are uncomfortable with being forced to increase activity. This underpins that the motive behind closet indexing is a lack of new quality bets.

### 6.2 An alternative test of decreasing return to scale?

One constraint discussed in the active management literature is decreasing returns to scale. If scale impacts performance, skill and scale interact, a more skilled large fund can underperform a less skilled small fund. Therefore, to learn about the skill, we must understand the effects of scale. The intervention we examine gives a unique opportunity to test whether skill has decreasing returns to scale. Suppose we choose to accept that asset managers are forced to increase their level of activeness. In that case, we have a quasi-natural experiment that allows us to investigate a sudden increase in active management. We assume managers cannot use in- or outflows to change the degree of active management. With a decreasing return to scale, we expect that a manager's new bets will be worse than the initial bets. Thus, if closet index funds underperform a comparable group of funds, then the marginal bets add less value than the initial bets, and we observe decreasing returns to scale.

When we analyze the total portfolio in Section 5.2, we find that closet index funds under scrutiny underperform relative to funds not under scrutiny. A more direct test of decreasing returns to scale is identifying where the new or the marginal bets are underperforming. In Table 6, in the previous section, we documented that the primary source of underperformance came from the new stock portfolio. In the performance evaluation of these sub-portfolios, we find significant losses for the funds forced to increase activity. This finding supports that returns to scale is negative for active management.

### 6.3 True active fund as control funds

As an alternative control group, we use another group of funds, not under scrutiny. This group consists of truly active funds in Scandinavia. However, these funds have differing characteristics. If we consider closet indexing to be a disease, contrasting closet index funds and truly active funds is tantamount to comparing sick funds with healthy ones. Such a comparison introduces a selection bias. Fund fixed effects may partly remedy this problem. As a robustness test of our main results, we carry out the same analyses as earlier.

In Table 7 we present the same summary statistics as we did in Table 2, but now the control group is truly active funds. There are three main differences between the samples. First, active share is lower for the closet index funds. Second, even if not much, about 24 basis points annually, the group of closet index funds is cheaper than truly active funds. Finally, the closet index funds are older than truly active funds. This confirms that closet index funds often can be old funds with uninformed investors.

**Table 7.**
**Summary statistics: Scandinavian funds**

This table presents summary statistics for the sample of closet index and truly active funds in Scandinavia (Denmark, Norway, and Sweden). The treated group are the Scandinavian funds classified as closet indexers, based on the active share limit highlighted in the top row. Conversely, the control group is the sample of truly active mutual funds in Scandinavia. Active share, assets under management, and fund age is as of December 31, 2013 before the event window. Expense ratio, gross and net alpha estimates are means over a two-year window before the event.

| Limit active share | 40% | | | 50% | | |
|---|---|---|---|---|---|---|
| Group | Treated | Control | Diff | Treated | Control | Diff |
| Number of funds | 46 | 104 | | 75 | 75 | |
| Active share (%) | 30.40 | 60.80 | $-30.40^{***}$ | 35.90 | 67.10 | $-31.20^{***}$ |
| Gross alpha (%) | 0.09 | 0.06 | 0.03 | 0.10 | 0.03 | 0.07 |
| Expense ratio (%) | 0.10 | 0.12 | $-0.02^{***}$ | 0.10 | 0.12 | $-0.02^{***}$ |
| Net alpha (%) | -0.01 | -0.06 | 0.05 | 0.00 | -0.09 | 0.09 |
| AUM (million USD) | 589 | 362 | 226 | 463 | 401 | 62 |
| Fund age (years) | 17.1 | 13.0 | $4.1^{***}$ | 16.0 | 12.5 | $3.5^{***}$ |

In Table 8 we present how scrutiny influence closet index funds relative to truly active funds. The Table has three panels, where the dependent variable is either active share, expense ratio or gross alpha.

**Table 8.**
**Robustness: Effect from intervention on main outcomes**

This table present tests on the effect the intervention had on our main outcome variables for closet index funds. In panel I, the dependent variable is active share, in panel II expense ratio and in panel III the dependent variable in gross alpha. We use 40% and 50% as cutoff points to classify funds as closet indexers. Our control group is comprised of the truly active Scandinavian funds. The model in Column 1-2 and 6-7, labeled Diff post, is a cross-sectional OLS model, where we compare the collapsed group-wise means before and after the event. The difference between the estimates are due to the underlying sample, where we in Column 1 and 6 use the actual sample, whilst in Column 2 and 7 we use a matched sample based on one-year pre-event average gross alpha, fund age and assets under management. DiD is a difference-in-differences model and FE are panel data regressions with both time and fund fixed effects. Standard errors are presented in parentheses. In Column 4 and 9 standard errors are clustered at the fund level, whilst in Column 5 and 10 they are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

**Panel I: Active share**

| | 40% | | | | | 50% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diff post | Diff post | DiD | FE | FE | Diff post | Diff post | DiD | FE | FE |
| Scrutiny | $-24.342^{***}$ | $-21.696^{***}$ | | | | $-26.944^{***}$ | $-24.518^{***}$ | | | |
| | (2.431) | (2.546) | | | | (1.877) | (1.919) | | | |
| Scrutiny · Post | | | 3.579 | $3.617^{**}$ | $3.902^{**}$ | | | 1.316 | 1.299 | $2.667^{*}$ |
| | | | (3.314) | (1.625) | (1.528) | | | (2.632) | (1.601) | (1.550) |
| Controls | No | No | No | No | Yes | No | No | No | No | Yes |
| Fund FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Time FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Sample | Actual | Match | Actual | Actual | Actual | Actual | Match | Actual | Actual | Actual |
| N | 150 | 78 | 300 | 7,288 | 6,908 | 150 | 124 | 300 | 7,288 | 6,908 |
| Adjusted $R^2$ | 0.400 | 0.482 | 0.468 | 0.892 | 0.902 | 0.579 | 0.569 | 0.605 | 0.890 | 0.901 |

**Panel II: Expense ratio**

| | 40% | | | | | 50% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diff post | Diff post | DiD | FE | FE | Diff post | Diff post | DiD | FE | FE |
| Scrutiny | $-0.033^{***}$ | $-0.030^{***}$ | | | | $-0.023^{***}$ | $-0.018^{***}$ | | | |
| | (0.006) | (0.008) | | | | (0.005) | (0.006) | | | |
| Scrutiny · Post | | | −0.007 | −0.005 | −0.004 | | | −0.003 | −0.002 | −0.001 |
| | | | (0.009) | (0.005) | (0.005) | | | (0.008) | (0.004) | (0.005) |
| Controls | No | No | No | No | Yes | No | No | No | No | Yes |
| Fund FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Time FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Sample | Actual | Match | Actual | Actual | Actual | Actual | Match | Actual | Actual | Actual |
| N | 148 | 78 | 294 | 7,140 | 6,908 | 148 | 124 | 294 | 7,140 | 6,908 |
| Adjusted $R^2$ | 0.190 | 0.154 | 0.131 | 0.756 | 0.775 | 0.103 | 0.066 | 0.075 | 0.756 | 0.774 |

**Panel III: Gross alpha**

| | 40% | | | | | 50% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diff post | Diff post | DiD | FE | FE | Diff post | Diff post | DiD | FE | FE |
| Scrutiny | $-0.104^{**}$ | $-0.098^{**}$ | | | | $-0.113^{***}$ | $-0.098^{**}$ | | | |
| | (0.043) | (0.041) | | | | (0.040) | (0.040) | | | |
| Scrutiny · Post | | | $-0.129^{*}$ | $-0.109^{*}$ | −0.148 | | | $-0.181^{***}$ | $-0.166^{**}$ | −0.115 |
| | | | (0.073) | (0.059) | (0.102) | | | (0.067) | (0.074) | (0.123) |
| Controls | No | No | No | No | Yes | No | No | No | No | Yes |
| Fund FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Time FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Sample | Actual | Match | Actual | Actual | Actual | Actual | Match | Actual | Actual | Actual |
| N | 148 | 78 | 294 | 10,197 | 9,945 | 148 | 124 | 294 | 10,197 | 9,945 |
| Adjusted $R^2$ | 0.032 | 0.057 | 0.011 | 0.100 | 0.106 | 0.046 | 0.038 | 0.022 | 0.101 | 0.106 |

Panel I is the same as Table 3, but the control group is now truly active funds. In Column 1-5, we present the result for the cutoff of an active share of 40%. In Column 1, we find that the closet index funds have an active share 24% lower than truly active funds after the intervention. If we match based on skill, size and fund age, we find that the difference decreases to 22%. In Column 3, we compare the difference between before and after for treated and control funds. We find that treated funds increase on average their active share by 4% more than control funds. In Columns 4 and 5, we perform regression analyses with fixed effects at both fund level and for each month. The interaction effect between after and funds under scrutiny is similar to the diff-in-diff estimate.

In Column 6-10, the cutoff point is 50%. We should expect smaller estimates since the degree of scrutiny probably is negatively related to active share. This is also what we find. The estimates are smaller and not always statistically significant. However, independent of the cutoff level, the funds under scrutiny have increased the activity level relative to non-scrutinized funds.

Panel II in Table 8 is the same as Table 4, but with truly active funds as controls. Again the results are not surprising. As shown in Column 1-2 and 3-4, the truly active funds are also cheaper after the intervention. From the diff-in-diff analyses or the fixed effects regressions, the estimates are small and not statistically significant. Thus, we find that funds under scrutiny do not lower their costs more than truly active funds.

Panel III is equivalent to Table 5 except for the control group. The results are surprisingly similar to what we estimated with the other control group. The differences lie in the magnitude of the estimates, which are somewhat higher. In Column 1-5, we present the results for the 40% limit, while we in Column 6-10 present the result for the 50% limit. In Columns 1 and 6, we find that the closet index funds have on average a 10- 11 basis points lower monthly excess return than truly active funds. If we convert the monthly returns to yearly returns, the difference is about $1 - 1.5\%$. If we match the Scandinavian closet index funds based on skill, size and fund age, we find that the difference is similar. In Columns 3 and 8, we compare the difference between before and after for treated and control funds. For the 40% cutoff, the diff-in-diff is about 13 basis points per month or 1.5% annually, while for the 50% cutoff, the diff-in-diff is about 18 basis points per month or 2% annually. The fixed effect regressions yield similar estimates as the diff-in-diff analysis. Notice that we in Column 4 and 9 cluster standard errors at the fund level while we in Column 5 and 10 cluster standard errors independently along the fund and date dimensions. In conclusion, we confirm that our previous results also hold for a different control group, which suffers from selection bias.

## 7 Policy implication and conclusion

The current trend is that people increasingly become their own money managers. The landscape in which they make decisions are becoming increasingly complicated, and many often lack the necessary information and knowledge to judge the quality of the products they purchase. This information asymmetry creates incentives for opportunistic behavior on the part of service providers, leading to market failures. The existence of closet indexing can be viewed as a realistic example of this failure. If investors were able to evaluate the quality of their fund managers' services, they would not choose to pay high fees for closet indexing. Clearly, there is a role for financial authorities to align the interest of financial intermediates and investors.

In this paper, we have tried to identify what we can learn from the interventions in the Scandinavian countries. We highlight the following four results. First, a potential closet index fund chooses to increase activity and not reduce fees when under suspicion from the supervisory authorities. Second, when closet index funds force themselves to become more active, they perform worse than the closet index funds not under scrutiny. Third, the funds under scrutiny take positions in new stocks when they increase activity. Finally, the funds under scrutiny are losing relative to unaffected funds on these new marginal bets. This finding lends support to the hypothesis about decreasing returns to scale.

Regulators have two main alternatives. They can either force funds to increase the active management level or to reduce fees. So far, there has been little research investigating the effect of this intervention. We show that when the funds have a choice, they increase their level of activity. We suggest two different reasons behind the increased activity behavior; either that managers are skilled and have new ideas but are afraid of losing assets under management, or they lack ideas and will harvest as much revenue as possible. We find that the managers that increase activity do not add much value to the investors. This finding supports a story where revenue harvesting managers do not have a plethora of new and good ideas.

So, what can regulators learn from these findings? One of the most important motives for closet indexing is that the managers do not have many additional investment ideas, and they are afraid of losing revenue by revealing that they are not as skilled as other fund managers. The best solution is not to force these funds to increase activity but to force them to update investor information and reduce the fees they charge their investors. Regulatory authorities should be careful forcing potential closet index funds to become more active. However, it is not that simple in practice because a regulator cannot stop managers from increasing activity and thereby defend their high fees. If managers choose to increase their activity and are not skilled, poor results should lead to lower assets under management. In the most severe cases, the funds will be forced out of business. However, this process is slow and will cause significant losses for uninformed investors on the way.

Taken together, a regulator should not force funds to be more active. Instead, their role should be to identify closet index funds. If funds want to keep a low degree of activity, this should be reflected by the fee levels and the information set. If the trend of investigating closet index funds is maintained, the hope is that this will discourage new funds from choosing such a strategy. In this setting, the DNB case from Norway illustrates that legal settlements can also be made. For closet index funds that will become a truly active fund, intermediaries and information providers such as Morningstar should pay special attention to their performance. This point also illustrates the benefit of having independent fund providers (Stoughton, Wu and Zechner (2011)) and the problem with "own brand" funds (Jenkinson, Martinez, Cookson and Jones (2020)). Of course, more information, such as presenting active share and tracking error, should also be available to investors.

Mutual funds are made up of trillions of dollars. The funds' investment decisions determine where a significant proportion of capital is allocated in the economy (Pástor, Stambaugh and Taylor (2020)). A recent trend is that more capital is allocated away from active and into index funds (Investment

Company Institute (2019)).[27] Eventually, the flow of funds from active to passive has to stop. The cost of active funds goes down, making active funds more attractive relative to passive funds. However, we always need to look after the content of the active funds. The debate about closet indexing gained momentum after the paper by Cremers and Petajisto (2009) and is still ongoing. The conviction of DNB in the Norwegian Supreme Court in February 2020 is a milestone in this debate. To our knowledge, this is the first sanction made by the court system. Only the future will tell how the closet index problem evolves from here.

[27]There are many reasons for this development. One reason is the poor historical performance of actively managed funds; another is increased competition between funds due to new technology. Finally, the difference between what some active funds have promised and what they have delivered can cause investors to lose confidence in active funds.

## References


Agarwal, V., Gay, G. D. and Ling, L. (2014), 'Window dressing in mutual funds', *The Review of Financial Studies* **27**(11), 3133–3170.

Agnew, J., Balduzzi, P. and Sunden, A. (2003), 'Portfolio choice and trading in a large 401 (k) plan', *American Economic Review* **93**(1), 193–215.

Authority, F. C. (2019), 'Closet trackers'. Available at https://www.fca.org.uk/firms/authorised-and-recognised-funds/closet-trackers, Last accessed on 2020-09-10.

Banegas, A., Gillen, B., Timmermann, A. and Wermers, R. (2013), 'The cross section of conditional mutual fund performance in european stock markets', *Journal of Financial economics* **108**(3), 699–726.

Berk, J. B. and Green, R. C. (2004), 'Mutual fund flows and performance in rational markets', *Journal of Political Economy* **112**(6), 1269–1295.

Berk, J. B. and van Binsbergen, J. H. (2015), 'Measuring skill in the mutual fund industry', *Journal of Financial Economics* **118**(1), 1–20.

Berk, J. B. and van Binsbergen, J. H. (2016), 'Active managers are skilled: On average, they add more than $3 million per year', *Journal of Portfolio Management* **42**(2), 131–139.

Bertrand, M., Duflo, E. and Mullainathan, S. (2004), 'How much should we trust differences-in-differences estimates?', *Quarterly Journal of Economics* **119**(1), 249–275.

Bilias, Y., Georgarakos, D. and Haliassos, M. (2010), 'Portfolio inertia and stock market fluctuations', *Journal of Money, Credit and Banking* **42**(4), 715–742.

Brands, S., Brown, S. J. and Gallagher, D. R. (2005), 'Portfolio concentration and investment manager performance', *International Review of Finance* **5**(3-4), 149–174.

Brown, D. C. and Davies, S. W. (2017), 'Moral hazard in active asset management', *Journal of Financial Economics* **125**(2), 311–325.

Campbell, J. Y. (2016), 'Restoring rational choice: The challenge of consumer financial regulation', *American Economic Review* **106**(5), 1–30.

Chen, J., Hong, H., Huang, M. and Kubik, J. D. (2004), 'Does fund size erode mutual fund performance? the role of liquidity and organization', *American Economic Review* **94**(5), 1276–1302.

Cremers, K. M. and Petajisto, A. (2009), 'How active is your fund manager? a new measure that predicts performance', *Review of Financial Studies* **22**(9), 3329–3365.

Cremers, K. and Quinn, C. (2016), 'Do mutual fund investors get what they pay for: Securities law and closet index funds', *Virginia Law and Business Review* **11**, 31.

Cremers, M., Ferreira, M. A., Matos, P. and Starks, L. (2016), 'Indexing and active fund management: International evidence', *Journal of Financial Economics* **120**(3), 539–560.

Cremers, M., Fulkerson, J. A. and Riley, T. B. (2020*a*), 'Active share and the predictability of portfolio performance: Out-of-sample evidence from separate accounts', *Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3574654* .

Cremers, M., Fulkerson, J. A. and Riley, T. B. (2020*b*), 'Benchmark discrepancies and mutual fund performance evaluation', *forthcoming Journal of Financial and Quantitative Analysis* .

Edelen, R. M., Evans, R. B. and Kadlec, G. B. (2007), 'Scale effects in mutual fund performance: The role of trading costs', *Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=951367* .

ESMA (2016), 'Supervisory work on potential closet index tracking'. Available at https://www.esma.europa.eu/document/esma-updates-supervisory-work-closet-index-tracking.

ESMA (2020), 'Closet indexing indicators and investor outcomes'. Available at https://www.esma.europa.eu/sites/default/files/library/esmawp-2020-2_closet_indexing.pdf.

Fama, E. F. and French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* .

Fama, E. F. and French, K. R. (1996), 'Multifactor explanations of asset pricing anomalies', *Journal of Finance* **51**(1), 55–84.

Ferreira, M. A., Keswani, A., Miguel, A. F. and Ramos, S. B. (2013), 'The determinants of mutual fund performance: A cross-country study', *Review of Finance* **17**(2), 483–525.

Ferreira, M. A., Matos, P. and Pires, P. (2018), 'Asset management within commercial banking groups: International evidence', *Journal of Finance* **73**(5), 2181–2227.

Financial Conduct Authority (2016), 'Economics for Effective Regulation'. Available at https://www.fca.org.uk/publication/occasional-papers/occassional-paper-13.pdf, Last accessed on 2020-09-10.

Financial Supervisory of Denmark (2013), 'Market developments 2013'. Available at https://www.dfsa.dk/~/media/Tal-og-fakta/2014/Markedsudvikling/Market-developments-2013-Investment-associations.pdf?la=en, Last accessed on 2020-09-10.

Financial Supervisory of Sweden (2015), 'Consumer protection on the financial market'. Available at https://www.fi.se/contentassets/9791b37d1d074ed4a3fc193786d449f6/konsumentrapp_2015engny.pdf, Last accessed on 2020-09-10.

Hitesh Doshi, R. E. and Simutin, M. (2015), 'Managerial activeness and mutual fund performance', *The Review of Asset Pricing Studies* **5**(2), 156–184.

Huang, J., Sialm, C. and Zhang, H. (2011), 'Risk shifting and mutual fund performance', *The Review of Financial Studies* **24**(8), 2575–2616.

Investment Company Institute (2019), 'A Review of Trends and Activities in the Investment Company Industry'. Available at https://ici.org/pdf/2019_factbook.pdf, Last accessed on 2020-09-10.

Jenkinson, T., Martinez, J., Cookson, G. and Jones, H. (2020), 'Best buys and own brands: Investment platforms' recommendations of mutual funds', *Review of Financial Studies* .

Kacperczyk, M., Nieuwerburgh, S. V. and Veldkamp, L. (2014), 'Time-varying fund manager skill', *The Journal of Finance* **69**(4), 1455–1484.

Kacperczyk, M., Sialm, C. and Zheng, L. (2005), 'On the industry concentration of actively managed equity mutual funds', *The Journal of Finance* **60**(4), 1983–2011.

Kjørven, M. E. (2019), 'Closet index funds and retail investor protection - a scandinavian perspective', *Tilburg Law Review* **24**(1), 125–138.

Leippold, M. and Rueegg, R. (2020), 'How rational and competitive is the market for mutual funds?', *Review of Finance* **24**(3), 579–613.

McLemore, P. (2019), 'Do mutual funds have decreasing returns to scale? evidence from fund mergers', *Journal of Financial and Quantitative Analysis* **54**(4), 1683–1711.

Miguel Anton, R. B. C. and Polk, C. (2020), 'Best ideas', *Available at https://papers.ssrn.com/sol3/papers.cfm?abid=1364827.* .

Morningstar (2019), 'U.S. Fund Fee Study'. Available at https://www.morningstar.com/articles/925303/2018-morningstar-fee-study-finds-that-fund-prices-continue-to-decline, Last accessed on 2020-09-10.

New York Office of the Attorney General (2018), 'Mutual Fund Fees and Active Share'. Available at https://ag.ny.gov/sites/default/files/ny_ag_report_on_mutual_fund_fees_and_active_share.pdf, Last accessed on 2020-09-10.

Norwegian Consumer Council (2015), 'Consumer Council sues Norway's largest bank'. Available at https://www.forbrukerradet.no/side/consumer-council-sues-norways-largest-bank-on-behalf-of-180000-consumers/, Last accessed on 2020-09-10.

Norwegian Ministry of Finance (2015), 'Financial Markets Report 2015'. Available at https://www.regjeringen.no/contentassets/fcbb121a9cb342f0b52d44300d3cc3b0/no/pdfs/stm201520160029000dddpdfs.pdf, Last accessed on 2020-09-10.

Pástor, L., Stambaugh, R. F. and Taylor, L. A. (2015), 'Scale and skill in active management', *Journal of Financial Economics* **116**(1), 23–45.

Pástor, L., Stambaugh, R. F. and Taylor, L. A. (2020), 'Fund tradeoffs', *forthcoming Journal of Financial Economics* .

Pastor, L. and Vorsatz, M. B. (2020), 'Mutual fund performance and flows during the covid-19 crisis', *forthcoming Review of Asset Pricing Studies* .

Petajisto, A. (2013), 'Active share and mutual fund performance', *Financial Analysts Journal* **69**(4), 73–93.

Phillips, B., Pukthuanthong, K. and Rau, P. R. (2018), 'Size does not matter: Diseconomies of scale in the mutual fund industry revisited', *Journal of Banking & Finance* **88**, 357–365.

Pollet, J. M. and Wilson, M. (2008), 'How does size affect mutual fund behavior?', *Journal of Finance* **63**(6), 2941–2969.

Pomorski, L. (2009), 'Acting on the most valuable information:'best idea' trades of mutual fund managers', *Available at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1108186* .

Reuter, J. and Zitzewitz, E. (2010), 'How much does size erode mutual fund performance? a regression discontinuity approach'. Available at https://www.nber.org/papers/w16329, Last accessed on 2020-09-10.

Sensoy, B. A. (2009), 'Performance evaluation and self-designated benchmark indexes in the mutual fund industry', *Journal of Financial Economics* **92**(1), 25–39.

Sirri, E. R. and Tufano, P. (1998), 'Costly search and mutual fund flows', *Journal of Finance* **53**(5), 1589–1622.

Stoughton, N., Wu, Y. and Zechner, J. (2011), 'Intermediated investment management', *Journal of Finance* **116**(3), 947–980.

Sunstein, C. R. (2015), 'Financial regulation and cost-benefit analysis', *The Yale Law Journal Forum* **124**, 263–279.

Swedish Ministry of Finance (2016), 'SOU 2016:45, En hållbar, transparent och konkurrenskraftig fondmarknad'. Available at https://www.regeringen.se/49f197/contentassets/c1e668bf3373497d88e7b0ef014d9f2a/en-hallbar-transparent-och-konkurrenskraftig-fondmarknad-sou-201645, Last accessed on 2020-09-10.

Yan, X. S. (2008), 'Liquidity, investment style, and the relation between fund size and fund performance', *Journal of Financial and Quantitative Analysis* **43**(3), 741–767.

Zhu, M. (2018), 'Informative fund size, managerial skill, and investor rationality', *Journal of Financial Economics* **130**, 114–134.

# Appendices

## A The DNB Norge case

On 21 June 2016, the Norwegian Consumer Council instituted legal proceedings before the Oslo District Court against DNB Asset Management AS (henceforth DNB), a wholly-owned subsidiary of DNB ASA offering asset management services. The Norwegian Consumer Council instituted a group action to pursue compensation on behalf of 180,000 investors in DNB Norge, a fund managed by DNB. The lawsuit alleged that the investor information and the high fees charged gave the investors the reason to expect active management. In contrast, the funds actually were managed very close to the index.

The Oslo District Court passed its judgment on 12 January 2018, whereby the claim was rejected, and DNB was held not liable. On 12 February 2018, the Norwegian Consumer Council appealed to the Borgarting Court of Appeal. The ruling from the Court of Appeal was announced on 8 May 2019 and ruled in favor of the Norwegian Consumer Council in the group action. [28]

In short, the Court of Appeal describes the effort of active management as (i) undertake analyses to identify potential good bets; and (ii) translate this into active positions such that the fund deviates from the index to the extent that is not insignificant. By comparing the investor information and the high fee on the one hand and the fact that the funds were managed very close to the index, the Court concluded that DNB had violated its obligations to its investors. DNB was sentenced to pay approximately NOK 350 mill (approx. USD 35 mill). [29]

DNB appealed the case to the Norwegian Supreme Court. The appeal case started on 21 January 2020. The ruling was delivered on 27 February, upholding the Court of Appeal's ruling.[30][31][32]

## B Data

### B.1 Sample selection and raw fund data

In this section, we present details on the sample selection and construction. Next, we present information on the fund data. This section is provided for information on how to replicate the results of the paper.

#### Sample selection

The initial sample is constructed based on lists generated in Morningstar Direct. As explained in the main part, and shown in Table 1, the treated countries are Denmark, Norway, and Sweden. In these countries, there has been strict scrutiny by the financial authorities. The European countries included in the initial sample are Austria, Belgium, Finland, Poland, Portugal, and Switzerland. In these countries, there has been no scrutiny or interference from the FSAs. For each country we construct lists based on the fields *Global Broad Category Group*, *Investment Area*, *Firm Country*, and *Base*

[28] Available at Lovdata TOSLO-2016-105341-2.
[29] Lovdata LB-2018-43087.
[30] Better Finance press release (2020).
[31] Press release Norwegian consumer council.
[32] Lovdata HR-2020-475-A.

*Currency*. We set the field *Global Broad Category Group* equal to equity to extract equity-only funds. Next we filter by *Firm Country* equal to *Investment Area* for funds with *Firm Country* for the countries included in the study, to obtain domestic funds. Last, we set the *Base Currency* equal to the domestic currency in each country. Also, a large part of the funds is structured with multiple share classes. We use the field *Oldest Share Class*, which takes the values of either Yes or No, to filter out the main share class of each fund. The initial sample consists of 1,148 funds, as presented in Table B.1.

**Table B.1.**
**Sample selection of domestic equity mutual funds**

This table presents the outcome of our sample selection procedure. The number of funds at the initial step are those where the management company is located within the same geographic area as they invest. At the fund type step we exclude all the funds that are registered as either an index fund, enhanced index fund or a fund of funds. To draw meaningful inference, we require that funds are alive one year before and after the event. Thus we exclude all funds that have an inception date after 31.01.2013 or an obsolete date before 31.12.2016 in the alive during event step. In order to form treatment and control groups we need data on active share before and after the event. Finally, we require funds to have data on key variables such as returns, size etc. during the event, and thus exclude funds that have missing observations over the two-year event period.

| Step | Total | Treated | Control |
|---|---|---|---|
| Initial | 1,148 | 624 | 524 |
| Fund type | 960 | 522 | 438 |
| Alive during event | 378 | 177 | 201 |
| Data coverage | 353 | 156 | 197 |
| **Total sample** | **353** | **156** | **197** |
| Active share < 50% | 122 | 47 | 75 |
| Active share < 40% | 79 | 33 | 46 |

Next, we impose three additional filters based on the fund type. As this study's scope is to interpret the portfolios of actively managed funds, we require the funds to be active, i.e., manage a portfolio where the objective is to outperform a passive benchmark index and have a managed portfolio. For this we use the fields: *Index Fund*, *Enhanced Index Fund*, and *Fund of Funds*. These three fields take the values Yes or No, and we set all these parameters to No. For robustness purposes, we cross-check the fields from Morningstar with the Lipper database and find that our initial sample selection is not free of errors. Despite having removed index and enhanced index funds from the sample before matching with Lipper, there are still three Swiss funds flagged by Lipper as index funds. With the two data providers categorizing the funds differently, we manually check the funds' investment objectives to determine which category is most appropriate. All of them state directly in the investment objective that they are either an index fund or replicate their benchmark index using either the physical or synthetic method. After excluding funds based on fund type, we are left with a sample consisting of 960 funds in total.

The last requirement is that the funds have sufficient data before and after the event window. The event is taking place between January 2014 and December 2015. To draw meaningful inference, we require that each fund has data starting, at the latest, one year prior and ending, at the earliest, one year after the event. We use the fields *Inception Date* and *Obsolete Date* to filter out funds. This means that funds with an inception date after January 2013 or an obsolete date before December 2016 are filtered out of the initial sample. This leaves us with a sample of 378 domestic actively managed equity mutual funds in our initial sample from which to draw treatment and control funds based on active share, with 177 potential treated funds and 201 potential control funds.

### Fund data

After defining the initial sample, we collect fund returns, fund size, and portfolio holdings. The main source of the time series fund data is Morningstar Direct, while we use both Morningstar and Lipper Database for the fund portfolio data. For each constituent in the lists explained in the previous section, we download the variables *Monthly Return*, *Monthly Gross Return*, and *Monthly Fund Size* aggregated over share classes.

For returns, all income and capital gains are reinvested monthly. The return data is in the local currency, while assets under management are in USD to have a common currency for comparison across countries. The Monthly Return includes management, administrative, and other costs that are deducted from the NAV, such as the 12b-1 fee, and gross returns are returns before fees. Thus, we use these two variables to compute the expense ratios following the definition from Morningstar Direct.

For the portfolio data, we use both the Morningstar and Lipper database. However, we find that some of the other European countries' funds are missing portfolio data in the Lipper database. For these funds, we download the portfolios from Morningstar to complete the data. We match the Morningstar (fund characteristics, performance, and portfolios) and Lipper data (fund portfolios) by ISIN or fund name if ISIN is missing. We end up with a link-list between the two databases with ISIN, fund names, Lipper IDs (Lipper's internal fund identifier) and Sec ID (Morningstar's internal fund identifier).

## B.2 Benchmark data

To measure active share and compare the fund returns to the returns of a benchmark, we must determine a benchmark index to which the fund portfolios and performance are evaluated. We use Datastream to download the constituents and benchmark weights, as well as the benchmark returns. We use the primary prospectus benchmark from Lipper if these are available. For funds where the primary prospectus benchmark constituents are unavailable in Datastream, we choose to use the most common domestic benchmark within each country for that particular fund type.[33] Also, for some of the indices, we cannot obtain the actual index weights from Datastream and use value-weighted weights based on market capitalization for the constituents. We match the benchmark portfolios with the fund portfolios based on stock ISIN. After downloading data for the initial fund sample, some funds miss either fund or portfolio data. After excluding funds with missing data, we end up with a sample consisting of 353 funds in total, where 156 are potential treated funds, and 197 are potential control funds.

The last row in Table B.1 reports the final sample. The treated funds are funds from Scandinavia, and the control funds are funds from other European countries. We also show how many funds are closet index funds based on a limit of an active share of either 50% or 40%.

[33] This applies in funds are divided into funds focusing on small, mid, or large-cap companies.

## B.3 Sample

Table B.2, reports summary statistics for all the domestic European active funds in the sample. The table presents the base sample of funds, where we form treatment groups in the main tests based on active share levels.

**Table B.2.**
**Summary statistics: Equity mutual funds**

This table presents summary statistics for the full sample of funds. The full sample of treated funds are all Scandinavian funds, and the full sample of control funds are all funds from other European countries. Active share, assets under management, and fund age is as of December 31, 2013 before the event window. Expense ratio, gross and net alpha estimates are means over a two-year window before the event.

| Sample | Full sample | Treated | Control | Diff |
|---|---|---|---|---|
| Number of funds | 353 | 156 | 197 | |
| Active share (%) | 55.6 | 51.5 | 59.7 | $-8.2^{***}$ |
| Gross alpha (%) | 0.07 | 0.06 | 0.09 | $-0.03$ |
| Expense ratio (%) | 0.11 | 0.11 | 0.11 | 0.00 |
| Net alpha (%) | -0.04 | -0.06 | -0.02 | $-0.04$ |
| AUM (million USD) | 295 | 423 | 193 | $229^{***}$ |
| Fund age (years) | 12.8 | 14.0 | 11.8 | $2.2^{**}$ |

## B.4 Return factors

We download style portfolio returns from MSCIs webpage and create factor returns according to Fama and French (1993, 1996) and the methodology described on Ken French's webpage.[34] As market return we use each countries well diversified all cap portfolio with neutral loading from MSCI. To create the size and value factor we download large cap (LC) and small cap (SC) returns, both with returns for value, none, and growth style indices to calculate the size and value factor. All returns are downloaded in local currency to correspond to the fund returns. The factor returns are then calculated by the following formulas:

$$\text{SMB} = \frac{1}{3}\left(SC_{value} + SC_{none} + SC_{growth}\right) - \frac{1}{3}\left(LC_{value} + LC_{none} + LC_{growth},\right) \tag{1}$$

$$\text{HML} = \frac{1}{2}\left(SC_{value} + LC_{value}\right) - \frac{1}{2}\left(SC_{growth} + LC_{growth}.\right) \tag{2}$$

where the subscripts denote the style of the portfolios. Using MSCI returns to create factors, we get homogeneity across the countries, as the returns of these factors are calculated in the same way for all countries.

[34] Available at: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/f-f_factors.html.

## B.5 Variables

In this section, we present the variables divided into outcome variables and control variables. The outcome variables are tested in the regressions, and the control variables are included in the vector of controls. We use control variables widely used in the literature.

**Table B.3.**
**Variable definitions**

This table documents our variables and their definitions. We group them into two categories: Outcome variables and explanatory variables.

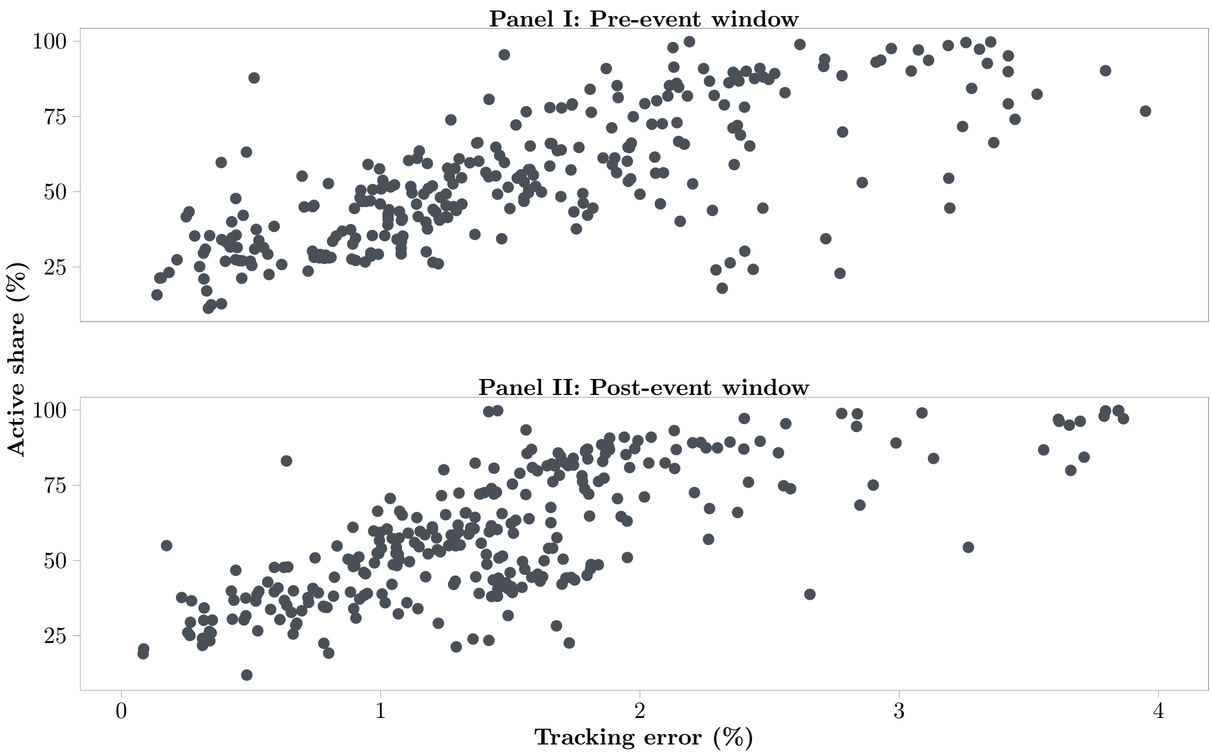
| **Outcome variables** | |
|---|---|
| Name | Definition |
| Active share | Percentage of a fund's portfolio holdings that differ from its benchmark index holdings. |
| Expense ratio | Monthly expense ratio. |
| Gross alpha | Difference between the fund gross return and its benchmark return. |
| Net alpha | Difference between the fund net return and its benchmark return. |
| Factor adjusted alpha | Three-factor alpha (percentage per month) with country-specific factors. |
| Gross value added | Product of a fund's gross alpha and size (AUM) in USDm. |
| **Control variables** | |
| Name | Definition |
| AUM | Total assets under management in USDm for all share classes. |
| Family AUM | Total assets under management in USDm for the funds in the same management company excluding the fund's own AUM. |
| Fund age | Number of years since the fund's launch date. |
| Flows | Percentage growth in AUM, net of internal growth (assuming reinvestment of dividends and distributions). |
| Explicit indexing (% AUM) | Percentage that explicitly indexed funds represent of the AUM of open-end equity mutual funds in the fund's country. |
| Fund industry size | Sum of total assets under management for open-end equity mutual funds in the fund's country. |
| GDP per capita | Gross domestic product per capita in U.S. dollars in the fund's country (World Development Indicators). |

## B.6 Relations between active share and tracking error

An alternative measure of active management level is tracking error, i.e., the standard deviation of the funds' active returns. Active share and tracking error are often used in combination to determine whether funds are potential closet indexers, where the active share is forward-looking while tracking error requires historical data for calculation. Figure B.1 plots the mean active share against tracking error of monthly return observations in the pre- and post-event window, in Panel I and II, respectively. The correlation coefficients between the two variables are 0.73 in the pre-event window and 0.74 in the post-event window. This shows that for the domestic funds in our sample, these two measures are highly correlated. This confirms the findings from ESMA (2020).

**Figure B.1.**
**Tracking error and active share**

This figure shows the relationship between tracking error and active share in the fund sample. Panel I plots it for the pre-event window, and panel II for the post-event window.




# C Additional analyses regarding impact on active share and fees

## C.1 Testing for pre-trends

If the groups have a differential trend in the pre-event window, this can be a concern. If the interaction between treatment and a linear time trend is significant and of the same sign as the corresponding interaction, the results might be a manifestation of the favorable trend rather than the effect of the intervention.

Figure 2 shows that there is no large difference in pre-trend for active share. In Table C.1 we formally test for pre-trends. For the cutoff point at 50% the pre-trends are close to zero. For the cutoff point at 40%, the trend is significantly negative without controls, while insignificant for the regression with controls. Notice that the pre-trend has an opposite sign of the corresponding interaction terms in Table 3. A concern with the trend with opposite signs is that the effect is based on reversion to the mean. However, given the magnitude of the coefficients in our regressions, we argue that the effect goes far beyond reversion to the mean.

Figure 3 shows that there is no large difference in pre-trend for expense ratio. In Table C.2 we formally test for pre-trends. The trend before the event is close to zero and insignificant for all cutoff points and regression specifications.

**Table C.1.**
**Testing for pre-trend in active share**

This table presents formal tests of pre-trend in the main regressions testing the effect the intervention had on active share in Table 3 in Section 5. The displayed results are from panel regression over the pre-event window. The regressions in Column 1-2 employs the sample of closet index funds using the 40% active share limit. Conversely, Column 4-5 uses the 50% active share limit. The coefficient Scrutiny · t tests the difference in trend in active share. All regressions include fund fixed effects and time fixed effects. Standard errors are presented in parentheses. In Column 1 and 3 standard errors are clustered at the fund level. In Column 2 and 4 standard errors are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

| | 40% | | 50% | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Scrutiny · t | $-0.142^{**}$ | $-0.169$ | $-0.021$ | 0.013 |
| | (0.062) | (0.124) | (0.053) | (0.075) |
| Controls | No | Yes | No | Yes |
| Fund FE | Yes | Yes | Yes | Yes |
| Time FE | Yes | Yes | Yes | Yes |
| Sample | Actual | Actual | Actual | Actual |
| N | 1,864 | 1,711 | 2,875 | 2,637 |
| Adjusted $R^2$ | 0.887 | 0.908 | 0.907 | 0.915 |

**Table C.2.**
**Testing for pre-trend in expense ratio**

This table presents formal tests of pre-trend in the main regressions testing the effect the intervention had on expense ratio in Table 4 in Section 5. The displayed results are from panel regression over the pre-event window. The regressions in Column 1-2 employs the sample of closet index funds using the 40% active share limit. Conversely, Column 4-5 uses the 50% active share limit. The coefficient Scrutiny · t tests the difference in trend in active share. All regressions include fund fixed effects and time fixed effects. Standard errors are presented in parentheses. In Column 1 and 3 standard errors are clustered at the fund level. In Column 2 and 4 standard errors are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

| | 40% | | 50% | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Scrutiny · t | 0.000 | 0.000 | $-0.000$ | $-0.000$ |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Controls | No | Yes | No | Yes |
| Fund FE | Yes | Yes | Yes | Yes |
| Time FE | Yes | Yes | Yes | Yes |
| Sample | Actual | Actual | Actual | Actual |
| N | 1,861 | 1,729 | 2,879 | 2,667 |
| Adjusted $R^2$ | 0.968 | 0.978 | 0.949 | 0.955 |

## C.2 Propensity score matching

This section explains how we match treated funds with control funds. A propensity score is the probability of treatment conditional on the matching variables. The matching variables are realized gross alpha, assets under management in USD and fund age. The propensity scores are calculated by running the regression.

$$Treat_i = \beta_1 \bar{\alpha}_{i,2013m12} + \beta_2 AUM_{i,2013m12} + \beta_3 Age_{i,2013m12}, \tag{3}$$

where assets under management and fund age are collected in December 2013, and gross alpha are the means over the last year, i.e., from January 2013 to December 2013. We match with ratio 1-1, i.e., each of the treated funds is assigned one control fund with similar properties along with the matching variables. Also, we add constraints on how large the propensity score differences can be for the funds to be a good match. We set the constraint, *Caliper*, which is the standard deviation of the propensity score within which to draw control units to 0.25. Also, we use average alpha as a matching variable in the case of multiple potential matches within each limit.

### C.3 Signal jamming

We examine whether the change in activity comes from "real bets" or whether the managers only added noise to their returns, also known as signal jamming (Brown and Davies (2017) and Cremers et al. (2020$a$)). As a measure for whether the fund randomly selects stocks or take more concentrated bets, we use the Stock Concentration Index ($SCI$). This is a stock-level version of the industry concentration index developed by Kacperczyk et al. (2005). The $SCI$ is calculated by summarizing the squared active weights,

$$SCI = \sum_{i=1}^{N} (w_i^a)^2, \tag{4}$$

for a total of $N$ stocks. The measure gives larger weight to larger active weights and suppresses smaller active weights. The measure is related to a portfolio level of the Herfindahl measure ($HHI$). The advantage of $SCI$ is that it "corrects" for the market weights.[35] By comparing the development of $SCI$ for treated closet index funds with control closet index funds, we learn how the managers implement the activity change. If they spread the bets (signal jamming), we would expect no difference in change of $SCI$, while if they take more concentrated bets, we would expect an increase in the difference in change of $SCI$.

To describe the development of the active portfolio concentration, we use the same methodology as when testing the change in active share. We compare the $SCI$ before and after the event and against a control group in a diff-in-diff setting.[36]

We report the regression results in Table C.3. Suppose we assume the active portfolios for treatment and control are equal before. In that case, we find from Column 1 and 2 that the treated post portfolios with a cutoff at 40% have more concentrated active bets. This is not the same for the 50% cutoff. In the diff-in-diff analysis and fixed effect regression for both cutoffs, the results are mixed. Thus, mixed results regarding how the managers implement the increased active share for the active portfolios. In Appendix 6.1, we perform similar analyses to investigate different parts of the active portfolios. Among other things, we find that treated managers take more concentrated bets in stocks not previously part of the active portfolio.

[35] The Herfindahl-Hirschman Index is defined as $HHI = \sum_{i=1}^{N} (w_i^p)^2$. The weight $w_i^p$ is total portfolio holding, and not the active weight $w_i^a$.

[36] The numbers for $SCI$ are presented in percentage. A fund with three active weights of $-0.05$, $-0.05$, and $0.10$ have $SCI$ equal to $(-0.05)^2 + (-0.05)^2 + (0.10)^2 = 0.015$, and this is presented as 1.5 (in %).

**Table C.3.**
**Effect from intervention on SCI**

This table present tests on the effect the intervention had on SCI for closet index funds. We use 40% and 50% as cutoff points to classify funds as closet indexers. The model in Column 1-2 and 6-7, labeled Diff post, is a cross-sectional OLS model, where we compare the collapsed group-wise means before and after the event. The difference between the estimates are due to the underlying sample, where we in Column 1 and 6 use the actual sample, whilst in Column 2 and 7 we use a matched sample based on one-year pre-event average gross alpha, fund age and assets under management. DiD is a difference-in-differences model and FE are panel data regressions with both time and fund fixed effects. Standard errors are presented in parentheses. In Column 4 and 9 standard errors are clustered at the fund level, whilst in Column 5 and 10 they are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: $^{***}$p<0.01, $^{**}$p<0.05, $^{*}$p<0.1.

| | 40% | | | | | 50% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diff post | Diff post | DiD | FE | FE | Diff post | Diff post | DiD | FE | FE |
| Scrutiny | 0.012$^{***}$ | 0.012$^{***}$ | | | | 0.000 | −0.000 | | | |
| | (0.002) | (0.002) | | | | (0.006) | (0.006) | | | |
| Scrutiny · Post | | | 0.006 | 0.006$^{***}$ | −0.018$^{***}$ | | | −0.001 | 0.003 | −0.026$^{**}$ |
| | | | (0.004) | (0.002) | (0.006) | | | (0.008) | (0.003) | (0.010) |
| Controls | No | No | No | No | Yes | No | No | No | No | Yes |
| Fund FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Time FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Sample | Actual | Match | Actual | Actual | Actual | Actual | Match | Actual | Actual | Actual |
| N | 79 | 77 | 158 | 3,788 | 3,691 | 122 | 119 | 244 | 5,697 | 5,506 |
| Adjusted $R^2$ | 0.236 | 0.233 | 0.110 | 0.764 | 0.815 | 0.000 | 0.000 | 0.001 | 0.412 | 0.421 |

# D Additional analyses regarding impact on value creation

## D.1 Testing for pre-trends

A different trend in the pre-event window between treated and control funds can be a concern. If the interaction between treatment and a linear time trend is significant and of the same sign as the corresponding interaction, the results might be a manifestation of the favorable trend rather than the effect of the intervention.

For gross alpha, Figure 4 shows that there is no large difference in pre-trend; if something, the treated funds perform better than the control funds. In Table D.1 we formally test for pre-trends. The trend before the event is close to zero and insignificant for all cutoff points and regression specifications.

**Table D.1.**
**Testing for pre-trend in gross alpha**

This table presents formal tests of pre-trend in the main regressions testing the effect the intervention had on gross alpha in Table 5 in Section 5. The displayed results are from panel regression over the pre-event window. The regressions in Column 1-2 employs the sample of closet index funds using the 40% active share limit. Conversely, Column 4-5 uses the 50% active share limit. The coefficient Scrutiny · t tests the difference in trend in active share. All regressions include fund fixed effects and time fixed effects. Standard errors are presented in parentheses. In Column 1 and 3 standard errors are clustered at the fund level. In Column 2 and 4 standard errors are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

| | 40% | | 50% | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Scrutiny · t | −0.002 | −0.006 | 0.004 | 0.010 |
| | (0.005) | (0.020) | (0.004) | (0.017) |
| Controls | No | Yes | No | Yes |
| Fund FE | Yes | Yes | Yes | Yes |
| Time FE | Yes | Yes | Yes | Yes |
| Sample | Actual | Actual | Actual | Actual |
| N | 1,859 | 1,729 | 2,876 | 2,667 |
| Adjusted $R^2$ | 0.078 | 0.084 | 0.121 | 0.117 |

## D.2 Clustered standard errors

A potential concern with the regressions in the main analysis is that the time dummy does not fully capture potential autocorrelation across funds each month. Therefore we extend the results in Table 5 in Section 5.2 by clustering on date and benchmark pairs. This clustering method implies that we allow for autocorrelation. From the different clusters in Table D.2, we find that the standard errors increase when we go from one-way fund clustering to two-way fund and time clustering.

**Table D.2.**
**Sensitivity: Cluster robust s.e. and gross alpha**

This table shows the effect of different clustering levels on the size of regression standard errors for the panel fixed effects regressions in Table 5, Column 5 and 10 in Section 5. Results are shown for the samples using 40% and 50% cutoffs. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

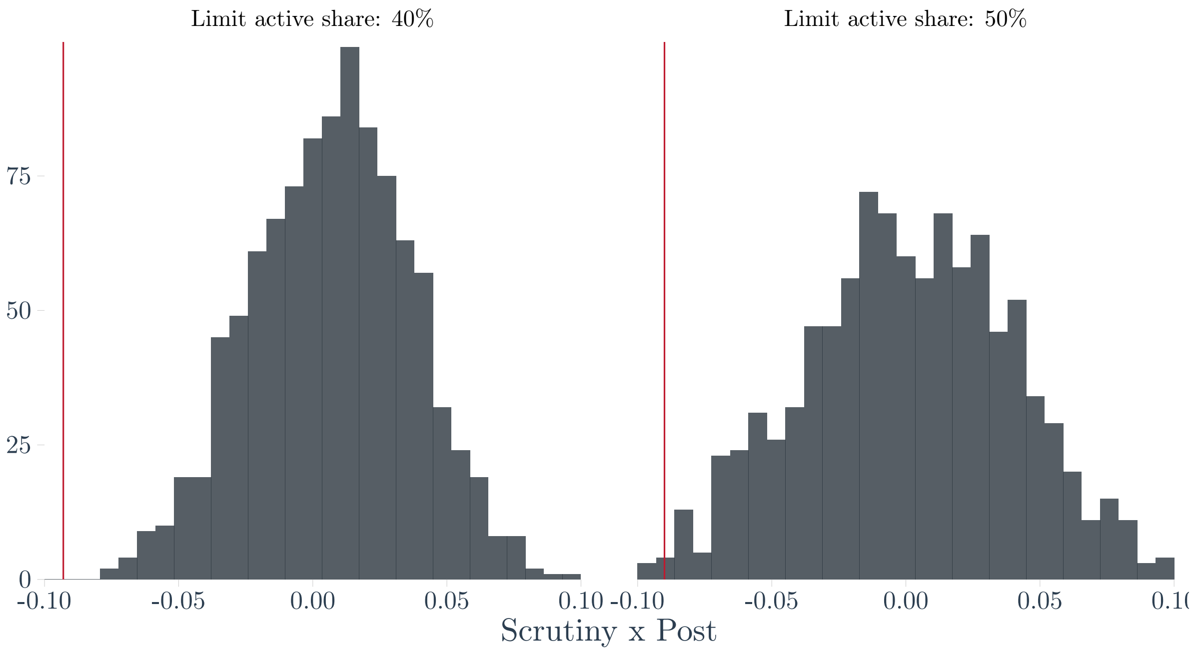
| | 40% | | | 50% | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Scrutiny · Post | −0.529*** | −0.529** | −0.529** | −0.396*** | −0.396* | −0.396 |
| | (0.086) | (0.257) | (0.265) | (0.090) | (0.221) | (0.241) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| FE | Fund + Time | Fund + Time | Fund + Time | Fund + Time | Fund + Time | Fund + Time |
| Cluster | Fund | Fund + Country x Time | Fund + Time | Fund | Fund + Country x Time | Fund + Time |
| N | 5,427 | 5,427 | 5,427 | 8,258 | 8,258 | 8,258 |
| Adjusted $R^2$ | 0.106 | 0.106 | 0.106 | 0.131 | 0.131 | 0.131 |

### D.3 Placebo tests

As shown in the histogram in Figure D.1, we find a significantly negative effect (5% confidence level) only for 1 of the 1,000 trials (0.1%) when using 40% as the limit on active share. For the sample using 50% as a limit, the corresponding number of trials with a significantly negative estimate is 13 (1.3%). Further, only 0.1% (40% limit) and 1.3% (50% limit) of the estimated coefficients on the Scrutiny · Post interaction term is equal to or smaller than the coefficient estimated using the original sample (-0.093 and -0.090, represented by solid vertical lines in the figure). These results reassure us that our tests capture the treatment effect of regulatory scrutiny on fund value creation and not some random effect or omitted variable.

**Figure D.1.**
**Random assignment to treatment and control group**

This figure presents histograms of the estimated coefficients of a falsification test for the difference-in-difference model for gross alpha. We use 40% and 50% as the cutoff points to classify funds as closet indexers. In each of the 1,000 separate estimations, the treatment and control groups are randomly assigned following a uniform distribution with the ratio of treated to control identical to the one of the original sample (see Table 2). Then, the DiD model (i.e., a model equivalent to Column 3 and 8 in Table 5) is re-estimated using the randomly assigned treatment variable. The reported coefficients are for the interaction Scrutiny · Post. The solid vertical lines marks the estimated coefficients in Column 3 and 8 in Table 5.



## D.4 Net alpha

In Table D.3 we perform the same analyses for net alpha as we perform for gross alpha in Table 5. The estimates are similar.

**Table D.3.**
**Effect from intervention on net alpha**

This table present tests on the effect the intervention had on net alpha for closet index funds. We use 40% and 50% as cutoff points to classify funds as closet indexers. The model in Column 1-2 and 6-7, labeled Diff post, is a cross-sectional OLS model, where we compare the collapsed group-wise means before and after the event. The difference between the estimates are due to the underlying sample, where we in Column 1 and 6 use the actual sample, whilst in Column 2 and 7 we use a matched sample based on one-year pre-event average gross alpha, fund age and assets under management. DiD is a difference-in-differences model and FE are panel data regressions with both time and fund fixed effects. Standard errors are presented in parentheses. In Column 4 and 9 standard errors are clustered at the fund level, whilst in Column 5 and 10 they are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

| | 40% | | | | | 50% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diff post | Diff post | DiD | FE | FE | Diff post | Diff post | DiD | FE | FE |
| Scrutiny | $-0.049^{*}$ | $-0.043$ | | | | $-0.077^{***}$ | $-0.077^{***}$ | | | |
| | (0.026) | (0.026) | | | | (0.029) | (0.029) | | | |
| Scrutiny · Post | | | $-0.085^{*}$ | $-0.076^{*}$ | $-0.528^{*}$ | | | $-0.087^{*}$ | $-0.073^{**}$ | $-0.396$ |
| | | | (0.046) | (0.040) | (0.265) | | | (0.046) | (0.037) | (0.242) |
| Controls | No | No | No | No | Yes | No | No | No | No | Yes |
| Fund FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Time FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Sample | Actual | Match | Actual | Actual | Actual | Actual | Match | Actual | Actual | Actual |
| N | 79 | 77 | 158 | 5,522 | 5,427 | 122 | 119 | 244 | 8,445 | 8,258 |
| Adjusted $R^2$ | 0.031 | 0.021 | 0.004 | 0.096 | 0.106 | 0.049 | 0.050 | 0.011 | 0.128 | 0.132 |

## D.5 Factor adjusted alpha

So far, our analysis has not disentangled the excess return in either stock picking or factor tilts. In this section, we adjust the active returns for factor risk from the market, a size factor, and a value factor. One problem with measuring factor tilts is that we measure the exposure ex post, and this exposure can be different from the indented ex ante exposure.

The risk-adjusted diff-in-diff estimates for alpha from Table D.4 are consistent with the findings in Section 5, but are slightly smaller in magnitude. Thus, the negative excess return is caused by the combination of poor stock picking and factor exposure.

**Table D.4.**
**Effect from intervention on factor adjusted alpha**

This table present tests on the effect the intervention had on factor adjusted alpha for closet index funds. We use 40% and 50% as cutoff points to classify funds as closet indexers. The model in Column 1-2 and 6-7, labeled Diff post, is a cross-sectional OLS model, where we compare the collapsed group-wise means before and after the event. The difference between the estimates are due to the underlying sample, where we in Column 1 and 6 use the actual sample, whilst in Column 2 and 7 we use a matched sample based on one-year pre-event average gross alpha, fund age and assets under management. DiD is a difference-in-differences model and FE are panel data regressions with both time and fund fixed effects. Standard errors are presented in parentheses. In Column 4 and 9 standard errors are clustered at the fund level, whilst in Column 5 and 10 they are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: $^{***}$p<0.01, $^{**}$p<0.05, $^{*}$p<0.1.

| | 40% | | | | | 50% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diff post | Diff post | DiD | FE | FE | Diff post | Diff post | DiD | FE | FE |
| Scrutiny | −0.025 | −0.014 | | | | −0.001 | 0.002 | | | |
| | (0.033) | (0.032) | | | | (0.028) | (0.027) | | | |
| Scrutiny · Post | | | −0.077 | −0.067* | −0.019 | | | −0.046 | −0.038 | −0.025 |
| | | | (0.052) | (0.040) | (0.206) | | | (0.044) | (0.035) | (0.162) |
| Controls | No | No | No | No | Yes | No | No | No | No | Yes |
| Fund FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Time FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Sample | Actual | Match | Actual | Actual | Actual | Actual | Match | Actual | Actual | Actual |
| N | 79 | 77 | 157 | 5,492 | 5,427 | 122 | 119 | 243 | 8,410 | 8,258 |
| Adjusted $R^2$ | 0.008 | 0.02 | 0.007 | 0.089 | 0.093 | 0.000 | 0.000 | 0.002 | 0.098 | 0.097 |

To sum up, our results support that the true active managers were the most successful managers both in their stock selection and factor tilts. We have not found evidence indicating that closet index managers are more skilled than true active managers.

## D.6 Value added

Following the arguments in Berk and van Binsbergen (2015) and Berk and van Binsbergen (2016) value added for a period (before fees) can be viewed as the total value the funds extract from the capital markets. This is the product of gross active return, $\alpha$, multiplied with the fund's assets under management, AUM, at the end of the previous period.

In Table D.5 we perform the same analyses as in Table 5, but now with value added as the dependent variable. Even if the statistical significance is slightly lower, we find that funds under scrutiny extract less value from the stock market than funds not under scrutiny. In Column 1, we find that with a cutoff point at 40% the treated funds have destroyed, on average, 0.10 million USD more value than the control funds per month for the four years after the event. From Column 6, we find that the treated funds with a cutoff at 50% have destroyed , on average, 0.07 million USD more than the control funds per month for the four years after the event. The value treated funds have destroyed is larger with a limit at 40%.

**Table D.5.**
**Effect from intervention on gross value added**

This table present tests on the effect the intervention had on gross value added for closet index funds. We use 40% and 50% as cutoff points to classify funds as closet indexers. The model in Column 1-2 and 6-7, labeled Diff post, is a cross-sectional OLS model, where we compare the collapsed group-wise means before and after the event. The difference between the estimates are due to the underlying sample, where we in Column 1 and 6 use the actual sample, whilst in Column 2 and 7 we use a matched sample based on one-year pre-event average gross alpha, fund age and assets under management. DiD is a difference-in-differences model and FE are panel data regressions with both time and fund fixed effects. Standard errors are presented in parentheses. In Column 4 and 9 standard errors are clustered at the fund level, whilst in Column 5 and 10 they are clustered independently along the fund and date dimensions (two-way clustering). Asterisks denote statistical significance: $^{***}$p<0.01, $^{**}$p<0.05, $^{*}$p<0.1.

| | 40% | | | | | 50% | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Diff post | Diff post | DiD | FE | FE | Diff post | Diff post | DiD | FE | FE |
| Scrutiny | −0.102 | −0.085 | | | | −0.066 | −0.057 | | | |
| | (0.131) | (0.133) | | | | (0.097) | (0.098) | | | |
| Scrutiny · Post | | | −0.356* | −0.326* | −0.593 | | | −0.078 | −0.045 | −0.412 |
| | | | (0.192) | (0.189) | (1.088) | | | (0.166) | (0.152) | (0.599) |
| Controls | No | No | No | No | Yes | No | No | No | No | Yes |
| Fund FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Time FE | No | No | No | Yes | Yes | No | No | No | Yes | Yes |
| Sample | Actual | Match | Actual | Actual | Actual | Actual | Match | Actual | Actual | Actual |
| N | 79 | 77 | 156 | 5,395 | 5,391 | 122 | 119 | 241 | 8,264 | 8,214 |
| Adjusted $R^2$ | 0.008 | 0.006 | 0.018 | 0.058 | 0.058 | 0.004 | 0.003 | 0.003 | 0.053 | 0.053 |

# Does Active Fee Predict Mutual Fund Flow? *
# - Price Sensitivity of Demand for Active Management

Trond Døskeland
Norwegian School of Economics

André Wattø Sjuve
Norwegian School of Economics

Andreas Ørpetveit
Norwegian School of Economics

**Abstract**

Active fee is the ratio between the excess cost of active management over the index alternative and the fund's activity level. We suggest a simple model that explains active capital allocations in the presence of time-varying active fee. We show that investors respond in accordance with the model's prediction of a negative relationship between active fee and subsequent flows. Our findings can be interpreted as negative price elasticity of demand for active management, where one standard deviation change in active fee translates into a shift in net flow of 83.4 bps. This effect implies a 42% change in the unconditional expected yearly flow. The time-series relation is driven by both the activity level and the excess cost of active management. The result also holds after controlling for Morningstar Ratings.

*First posted on SSRN: June 1, 2021. Send correspondence to André Wattø Sjuve, Norwegian School of Economics, Email: andre.sjuve@nhh.no and Andreas Ørpetveit, Norwegian School of Economics, Email: andreas.orpetveit@nhh.no.

# 1 Introduction

Despite the growing popularity of passive investing, a large part of mutual fund money is actively managed. An extensive literature examines what matters to active mutual fund investors. Many papers examine how investors respond to signals such as past performance and fees.[1] In this paper, we argue that *active fee* is a valuable signal for a fund's potential to beat its benchmark index. Active fee was first proposed by Cremers and Curtis (2016) and is the ratio between the excess cost of active management over the index alternative and the activity level of the fund. The measure can be interpreted as the unit price of active management. The lower the measure, the larger the upside potential of the active mutual fund. We examine whether investors agree with us and use active fee as a criterion to choose active mutual funds.

There is a sizable discussion on how mutual fund investors should select among active funds. One part of the literature examines whether it is possible to identify skilled funds[2], another part, to which this paper contributes, examine what investors actually do. What signals influence investors are not limited to efficiency and rationality but can reflect motivations such as advertising and salesmanship.[3] Recently there has been a discussion on whether mutual fund investors behave rationally and use a full asset pricing model when deciding which funds to buy and sell (see Barber, Huang and Odean (2016) (BHO) and Berk and van Binsbergen (2016) (BvB)) or react to different attention-grabbing or easy-to-process signals (for example Ben-David et al. (2019) and Evans and Sun (2021) show that Morningstar ratings drive fund flows).

The hypotheses we develop are based on something "in the middle". We suggest investors follow a simple signal in the form of active fee when picking active funds. This signal is not a result of either a complicated asset pricing model or an estimated prediction that can change due to "new regimes." Active fee can be seen as the unit price of active management. The numerator, the excess fee, is how much the given amount costs, and the denominator, the activity level, is how many units of active management the investor gets.[4] In Section 2 we formalize the idea by developing a simple model of fund flow in the presence of time-varying active fee. Our model's key implication is a negative time-series relation between active fee and subsequent fund flow. The result can be interpreted as negative price elasticity of demand for active management. Furthermore, we develop hypotheses that indicate that the response is negative to excess fees and positive for activity level if the skill level, measured with IR, is larger than the risk aversion parameter against active management.

As the mapping from behavior to beliefs is not one-to-one, investors' true beliefs are unknown. But in practice, we show that they utilize active fee to pick active funds and that this is a

[1] See Ben-David, Li, Rossi and Song (2019) for a comprehensive review.

[2] See Cremers, Fulkerson and Riley (2019) for a comprehensive review.

[3] See Gallaher, Kaniel and Starks (2006) and Roussanov, Ruan and Wei (2021) for evidence on the importance of marketing to attract flows for mutual funds.

[4] With unit pricing, we take the price of a product, for example, 1.59 for a 24-ounce jar of spaghetti sauce, and divides the total cost by a standard unit of measurement (such as ounces). We then have a simple price comparison point, i.e., 1.59 divided by 24 ounces = 0.07 per ounce. We can then use that unit price to compare whether we get the best deal. For example, what is the unit price of a 45-ounce jar of sauce at 3.69? When we do the math, 3.69 divided by 45 ounces = 0.08 per ounce, we find that the smaller jar is the best buy.

sophisticated response to a challenging problem. Consistent with the model, we find that a fund's active fee negatively predicts the fund's subsequent flow. For example, a one standard deviation increase in a fund's active fee translates to a decrease in annual net flow of 83.4 bps ($0.834\% = -2.98 \cdot 0.28$). This number is economically significant in that it involves a 42% ($= 0.834/1.99$) change in the unconditional expected yearly flow. We find that the negative effect comes from both the numerator, excess fee, and the denominator, activity level. The size of the effect is about the same when we control for variables known to influence fund flows, such as Morningstar ratings and past performance. Finally, the effect is less important for large funds.

We give new insight into what drives mutual fund flows. As noted earlier, this literature is extensive, and our contribution is to come up with a rational but straightforward measure with as little as possible estimation error. Investors respond to different simple signals, for example, Morningstar ratings. These measures are based on historical performance one hopes will last into the future. Active fee is almost as simple as the Morningstar stars, and all information is freely available, but often not shown in investor information. However, we suggest that increased awareness may lead to better investment decisions and greater competition in the industry.

Mutual fund flows help us understand the important influences on consumers' economic thinking since the decisions are the aggregate of a large portion of society. Since mutual fund shares are issued and redeemed at a fixed price, essentially no matter the quantity, we can study the behavior of investors without considering the counterparty to the trade, the mutual fund. Such an opportunity to explore one side of the market rarely exists elsewhere.

Given investors' ability to switch among funds and a large number of active mutual funds, almost all funds face multiple close substitutes. Thus, the opportunity for investors to choose among rival funds is well established. However, empirical evidence suggests there is significant price dispersion even when products are homogeneous. Carlin (2009) show that complexity increases the market power because it prevents some consumers from becoming knowledgeable about prices in the market.

A relevant question is whether investors are sufficiently sensitive to price that they will respond to excess fees by substituting into a reasonably close substitute fund with a lower price. The fund flow response to changes in active fee can be interpreted as an estimate of the degree of price sensitivity of demand for active equity mutual funds. Sirri and Tufano (1998) and Barber, Odean and Zheng (2005) find a negative relationship between total fund fee and mutual fund flow. However, to our knowledge, no other papers have examined the price elasticity of demand for active management with such a finely grained measure we use.

The rest of the paper is organized as follows: In Section 2, we propose a simple model and develop the hypothesis. In Section 3, we present the research design, including the data and sample. Section 4-6 shows the results, where the main result is described in Section 4. Finally, in Section 7, we conclude.

## 2 Model of the active fee-flow relation

This section outlines a simple partial-equilibrium ex-ante model of optimal active fee for an active mutual fund investor. Active fee is the ratio between the excess cost for active management over the degree of active management performed by the manager. The model suggests a negative active fee relation: a time-series regression in which a fund's active fee is negatively related to the fund's subsequent value-added. For the components of active fee, the activity level and excess fee, the relation is negative for fees and, in most cases, positive for the activity level.

Based on the pioneering work of Markowitz (1952), investors cannot expect to be rewarded, on average, for taking unsystematic risk. However, if an investor possesses the skill to select superior stocks, the optimal portfolio strategy for an investor can change. Forming a portfolio that concentrates on stocks for which the investor has the highest conviction may represent a dominant approach (see for example Merton (1987), and Nieuwerburgh and Veldkamp (2010)). Thus, a prospective mutual fund investor faces the choice of an index fund in line with Markowitz or an active fund that invests in roughly the same universe of assets as the index fund but with aspirations for better performance and the certainty of higher risk.

### Active fund selection model

Investors buy actively managed funds to earn excess returns, i.e., maximize risk-adjusted excess return net of fees. We call the difference in return between the managed portfolio and the passive index benchmark portfolio gross alpha, $\alpha_t^g$. A sensible heuristic that narrows the set of alternative passive assets is to focus on salient benchmarks. Such a benchmark model is relevant since there is no consensus regarding the true asset pricing model. Given model uncertainty, Ľuboš Pástor and Stambaugh (2002) show that investors can utilize passive assets (including benchmarks) instead of the unknown risk factors.

The risk associated with deviation from the benchmark is called activity level, $AL_t$. We set a negative weight on the deviation from the benchmark but opens up for that this parameter is low since an active manager can only add value relative to the index by deviating from it.[5] We name this active risk aversion parameter, $\lambda$.

The excess fee is the difference between the total price of the fund over the alternative index cost, $F_t$. As in Brown and Davies (2017), we split the fee into a part for portfolio management services (mainly diversification) and a part for active asset selection. Our fee measure is for the latter.[6] Thus, the investor trade off expected alpha against active risk and fees (Grinold and Kahn, 2000).[7] We assume the investors maximize the value-added, $VA_{t+1}$ given by the following equation:[8]

[5] Grinold and Kahn (2000), chapter 4, suggests that investors have two separate risk aversions, one for market risk and another for active risk. Since our model does not see the investment in a relationship with the investor's total portfolio, an extension of the model could incorporate the investor-specific diversification effect in this parameter.

[6] We are ignoring benchmark timing, assuming beta against benchmark equal to 1.

[7] We do not model the objective of the fund manager.

[8] Note that this is a different value-added measure than in Berk and Van Binsbergen (2015).

$$VA_{t+1} = \alpha_t^g - \lambda \cdot AL_t - F_t. \tag{1}$$

A fund manager who merely forms a concentrated portfolio without possessing the inherent stock selection skills is not likely to experience a successful result. As a measure of the consistency of performance, i.e., the probability that the manager will realize positive active return every period, we define information ratio $IR_t$. It helps little for the investor with a skilled manager (with positive IR) if the manager does not use his skill.[9] The way the manager can utilize his skills is by taking deviating positions from the index. The more and larger positions, the greater the activity level, $AL_t$. With a positive skill, this leads to larger gross alpha, i.e., $\alpha_t^g = IR \cdot AL_t$. Brown, Tiu and Yoeli (2020) also suggest a similar model with a positive relation between activity level and performance for skilled managers:

$$\begin{aligned} VA_{t+1} &= IR \cdot AL_t - \lambda \cdot AL_t - F_t \\ &= (IR - \lambda) \cdot AL_t - F_t. \end{aligned} \tag{2}$$

In this basic version of the model, we do not assume we can predict skill from previous signals. We acknowledge that identifying managerial skill is difficult given finite sample constraints; thus, for example, past performance or fees are not a signal for future success.[10] However, in some of the sub hypotheses, we will examine whether investors behave differently depending on the most common signals.

What will happen with an increasing level of activity or total fees? We first take the derivative of value-added with respect to the fee level,

$$\frac{\delta VA}{\delta F} = -1. \tag{3}$$

Since the derivative is negative, we want a lower fee. Then we take the derivative of value-added with respect to activity level:

$$\frac{\delta VA}{\delta AL} = IR - \lambda. \tag{4}$$

Our desired activity level will increase if the skill is more considerable than active risk aversion,i.e., $IR > \lambda$. Since a prerequisite of investing in active management is a belief in skill, $IR > 0$, and, as suggested above, $\lambda$ is low, we postulate a positive relation between $VA$ and $AL$. Furthermore, we find that the derivative is more positive for more skilled managers.

Thus, given that active fee is the ratio between excess fees and activity of the fund, i.e., $AF_t = \frac{F_t}{AL_t}$, a lower fee or a higher activity level first leads to a lower active fee, which again leads to a higher value-added for the investor.

[9] It does not help to have a skilled football player on the team if he sits on the bench for large parts of the match.
[10] Again see Cremers et al. (2019) for a review.

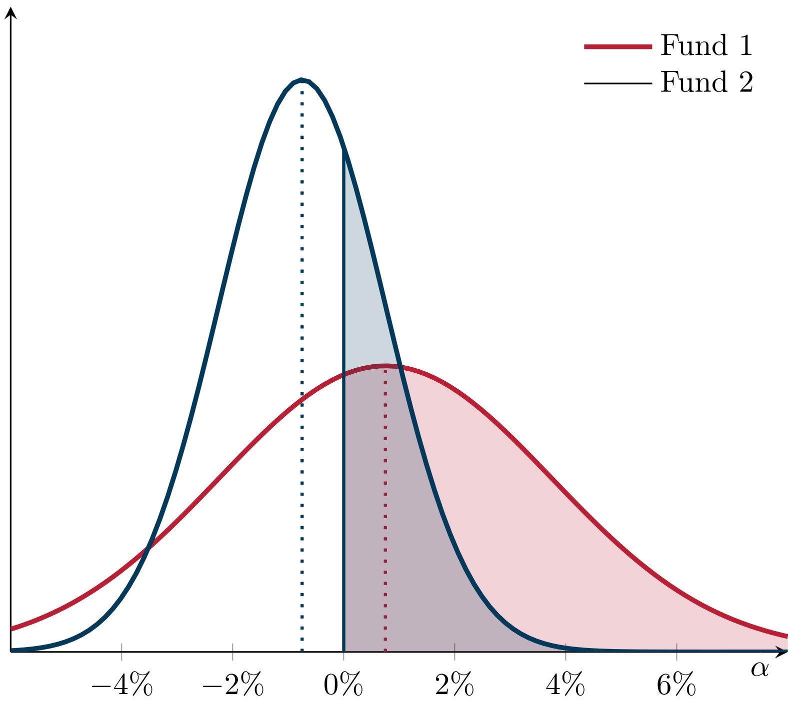
### Example: Comparing two funds

Before we formulate the hypotheses, we illustrate our simple model with an example. We compare two active funds, Fund 1 and Fund 2. The active fund literature has recognized that it is hard to predict specific estimates. For example, skill is random, leading us to assume skill, represented by $IR$, is the same for the two funds. Furthermore, we assume the risk aversion against active risk is identical. Thus, in our setup, activity level and fee are the only variables that differ between the funds, leading to a difference in value-added as follows:

$$VA_{1,t+1} - VA_{2,t+1} = IR \cdot (AL_{1,t} - AL_{2,t}) - (F_{1,t} - F_{2,t}). \tag{5}$$

**Figure 1:**
**Active return**

This plot illustrates the net alpha return distribution of two hypothetical funds with identical IR of 0.5, but different fee levels.




The net return of the two funds is shown in Figure 1. Fund 1 has an activity level, represented here by tracking error, at $\sigma(\alpha_1) = 3.0\%$, while manager 2 has a tracking error at $\sigma(\alpha_2) = 1.5\%$. We assume that both managers are equally skilled with an IR at 0.5. Fund 1 has an excess fee at 0.75%, while fund 2 has an excess fee at 1.5%. The red curve illustrates the investor's net alpha for Fund 1 where the expected return is $1.5\% - 0.75\% = 0.75\%$. The blue curve illustrated the distribution of net alpha for Fund 2 with expected return equal $0.75\% - 1.5\% = -0.75\%$. Using Equation 5, we find that the difference in value-added is 1.5%. Active fee is for fund 1: $AF_{1,t} = \frac{1.5\%}{1.5\%} = 1$, while active fee for fund 2 is $AF_{2,t} = \frac{0.75\%}{3.0\%} = 0.25$. This illustrate that a lower active fee leads to increased value-added.[11]

## Hypotheses

Our model's key implication is a negative time-series relation between active fee and subsequent fund flow. Below we list the main hypothesis and explain the sub hypotheses.

### Hypothesis 1: Active fee and subsequent fund flow

Above, we have argued that there should be a negative relation between active fee and subsequent fund flow, named H1a. The fund flow response to changes in active fee can be interpreted as an estimate of the degree of price sensitivity of demand for active equity mutual funds. In Equation 3 we showed that the relation between the excess fee level and subsequent fund flow is negative, named H1b. In Equation 4 we show that the association between activity level and subsequent fund flow depends on $IR$ and $\lambda$. Since a prerequisite of investing in active management is a belief in skill, $IR > 0$ and a low $\lambda$, we predict a positive relation, i.e., $IR > \lambda$, for the typical investor in H1c.[12]

- H1a: Negative relationship between active fee and subsequent fund flow.
- H1b: Negative relationship between excess fee and subsequent fund flow.
- H1c: Positive relationship between active share and subsequent fund flow.

### Hypothesis 2: Morningstar ratings and the active fee-fund flow relation

As also argued in the introduction, Morningstar rating has an important influence on fund flows. Mutual fund ratings do not produce new information. They are purely quantitative, backward-looking measures of funds' past performance (Huang, Li and Weng, 2020). In our simple model, we do not assume a relationship between past performance and fund flow. An interesting question is whether Morningstar ratings pick up the same signal as active fee. No paper we are aware of examines the relationship between past performance and active fee. However, Cooper, Halling and Yang (2020) suggest that there is no clear relationship. We examine how Morningstar ratings affect the relationship between active fee and fund flow in two steps. First, in H2a, we test if active fee is correlated with ratings, and then, in H2b, we test whether the impact from active fee changes when we control for the stars.

The degree of financial sophistication may also vary with the level of rating. Müller and Weber (2010) find that investors with high financial literacy choose funds by chasing fund performance.

[11]We can learn from Figure 1 that Fund 1 has a higher probability of net alpha relative to an index fund (the red shaded area) compare to Fund 2 (the blue shaded area). We also see that Fund 1 relative to Fund 2 has a larger upside potential, defined as probability of return above zero), however note that the likelihood of large losses are larger for Fund 1 than 2.

[12]Cremers, Ferreira, Matos and Starks (2016) find that funds with higher active share attract more flows. Sirri and Tufano (1998) and Barber et al. (2005) find a negative relationship between total fund fee and mutual fund flow.

This finding implies that owners of high-rated funds are more sophisticated than low-rated fund owners. A natural question then is whether financial sophistication is related to the response of active fee. We predict that more financial sophistication investors (high rating) respond more than less sophisticated investors (low ratings). This prediction we test in H2c.

- H2a: No relation between active fee and subsequent Morningstar rating.
- H2b: Negative relationship between active fee and subsequent fund flow after controlling for Morningstar rating.
- H2c: More negative relationship between active fee and subsequent fund flow for high rating investors.

**Hypothesis 3: Fund characteristics and the active fee-fund flow relation**

The effect may also vary for different segments of funds. First, in H3a, we examine how past performance interferes with the relation between active fee and fund flow. Past performance is easily accessible for investors and has been shown to be a predictor of subsequent flows (see for example Ferreira, Keswani, Miguel and Ramos (2012)). Since Morningstar rating builds on past performance, we suggest that past performance should have the same impact as Morningstar ratings (see H2c). Thus, we predict a more negative (positive) relationship between active fee and subsequent fund flow for funds with high (low) past performance.

In its simplest form, see Equation 2, we do not assume funds face decreasing returns to scale. However, suppose we expand the model to incorporate diminishing returns to scale in active management, as suggested by Berk and Green (2004). In that case, we find that skill, represented by $IR$, is lower for larger funds than for smaller funds. This again leads to a less negative relationship between active fee and fund flow. It is more important that a fund is active when skill is high than when skill is low. In H3b, we predict that the relationship between active fee and fund flow is more prominent for small funds than large funds.

Third, in H3c, we examine if the fund style interferes with the relationship between active fee and fund flow. As far as we find, little is known about the typical factor investor. Betermier, Calvet and Sodini (2017) find that value investors have higher financial and real estate wealth, lower leverage, lower income risk, lower human capital, and are more likely to be female than the average growth investor. However, since this is not much to back a hypothesis on, we are agnostic and formulate a null hypothesis that fund factors do not influence how active fee relates to fund flow.

- H3a: More negative relationship between active fee and subsequent fund flow for funds with high past performance.
- H3b: More negative (positive) relationship between active fee and subsequent fund flow for small (large) funds.

- H3c: Fama-French factors do not interfere with the relationship between active fee and subsequent fund flow.

# 3 Research design

This section describes the universe in which we test our hypotheses by detailing our data collection and sample construction. Furthermore, we discuss the main variables and offer some descriptive statistics.

## 3.1 Data and sample

We use Lipper as our primary database. It provides a comprehensive sample of mutual funds offered across many countries. Our primary interest is on the relation between active fee and flow; therefore, we focus on actively managed open-end equity mutual funds. The active fee calculation requires an activity measure. We use active share, introduced by Cremers and Petajisto (2009), which was known a few years before the paper's publication. Consequently, our sample period is 2006-2019.

From Lipper, we obtain fund characteristics (fund name, domicile, benchmark, monthly returns, total net assets (TNA), fees and expenses (TER)) in addition to detailed fund holdings. Since the database provides both active, merged and liquidated funds, it is survivorship bias-free. Lipper treats separate share classes within the same fund as different observations despite them having the same holdings and the same returns before expenses. Following Cremers et al. (2016), we keep as our unit of observation the share class that Lipper identifies as the primary share class. Fund-level variables are consequently aggregated across the different share classes. Fund holdings are necessary to calculate active share, but their availability varies across countries. In calculating active share, we follow Cremers et al. (2016) and use the Lipper-assigned benchmarks instead of the self-declared fund benchmarks to infer the investment opportunity sets of funds. In so doing, we avoid the problem of funds strategically picking their benchmarks. To arrive at securities' benchmark weights, we use the weights of index funds and ETFs that replicate the benchmark. Moreover, fund expense ratios are yearly fees. As highlighted in Coval and Stafford (2007), investors often respond slowly to fund performance. We, therefore, follow BHO and model the response of flows on a yearly basis.

The sample obtained from Lipper consists of 39,324 mutual funds with a combined TNA of 6.7 USDtn as of December 2019. In defining a fund's nationality, we follow Schumacher (2018), where it is the location of the management company that dictates residence over the fund's legal domicile. To estimate the relationship between flows and active fee, we require a fund to have at least 12 observations on size and returns. To compute active fee, we drop funds without any fee information or that have zero observations on active share. Given our primary interest in open-end equity mutual funds, we exclude index funds, fund-of-funds and closed-end funds. The resulting sample consists of 13,367 funds (122,172 fund-year observations) with 3.1 USDtn as of December 2019. Table 1 list the 30 sample countries along with the number of funds and fund-year pairs, assets under management, average net flow, active and excess fee and active share for each domicile.

To complement the data from Lipper, we gather additional information from Morningstar Direct on fund rating and Morningstar's style-box classification. Morningstar's Overall Rating is a time-series variable that assigns a fund a star rating between one and five stars on a monthly basis.[13] Style-box classification is a 3x3 grid variable indicating where a fund lies in the dimensions of size (small, mid, large) and style (value, growth, blend).[14] We match the Morningstar data into the Lipper data by ISIN or fund name (text-matching) if ISIN is missing.

## 3.2 Main variables

In Table A.1 in Appendix A.1 we give a complete list of all variables including their definition and data sources. Here, we give a brief overview of the main variables.

### Net flow

Following Chevalier and Ellison (1997) and Sirri and Tufano (1998), we define new money growth rate as the net growth rate in total net assets (TNA), not due to dividends and capital gains on the assets under management, but to new external money. Net flow (NF) for fund $i$ at time $t$ is calculated as:

$$\mathrm{NF}_{i,t} = \frac{\mathrm{TNA}_{i,t}}{\mathrm{TNA}_{i,t-1}} - (1 + r_{i,t}), \tag{6}$$

where $\mathrm{TNA}_{i,t}$ is the total net asset value in USD for fund $i$ at the end of time $t$ and $r_{i,t}$ is fund $i$'s total return from $t-1$ to $t$. Equation 6 assumes that flows occur at the end of each period, as we have no information on the timing of new investments. Since we have monthly information on fund size and returns, we compute net flows using Equation 6 monthly and aggregate these to a yearly variable, thereby avoiding the assumption that all flows occur at the end of each year. To ensure that extreme values do not drive our results, we winsorize fund flows at the bottom and top 1 % level of the distribution within each country-year.

### Active fee

An active fund's portfolio can be divided into an index fund and a "residual" fund with over- (long) or underweights (short). The relative weighting of these two parts is determined by the fund's activity level, for which there are several different measures.[15] We use active share (AS),

[13]See Ben-David et al. (2019) and Morningstar Ratings for details on Morgninstar ratings and their methodology.
[14]See Morningstar Style-Box for details on the Morningstar style-box.
[15]Three common measures are tracking error, active share (Cremers and Petajisto, 2009) and activeness (Pástor, Stambaugh and Taylor, 2020). With our forward-looking model, we prefer to use active share over ex-post tracking error. The information needed to compute active share is easier to obtain than ex-ante tracking error as well. Active share is also more readily available than the activeness measure by Pastor. However, there is a clear overlap between the different measures. For a formal analysis, see Marmoiton (2021).

as introduced by Cremers and Petajisto (2009).[16]

We define active fee (AF) as the unit price of the active management of the fund. Cremers and Curtis (2016) show that we can split a fund's total expense ratio into an index and an active part:

$$\frac{\text{TER} - (1 - \text{AS}) \cdot \text{Index fee}}{\text{AS}} = \frac{\text{TER} - \text{Index fee}}{\text{AS}} + \text{Index fee} \tag{7}$$

where we, in the second part, rewrite the expression to highlight the active and index fund part explicitly. The first term is the cost per percentage point of active share (activity level), i.e., what we name the active fee. The latter term is the cost associated with investing in an index fund tracking the same benchmark as the overall fund. In our analyses, we focus on the active fund component defined as:

$$\text{AF}_{i,t} = \frac{\text{TER}_{i,t} - \text{Index fee}_t}{\text{AS}_{i,t}} \tag{8}$$

The $TER$ term in Equation 8 is downloaded from Lipper and is the yearly total expense ratio of the fund. To obtain relevant index fees, we use our data on 3,100 index mutual funds and ETFs to compute benchmark-year fees for the benchmarks these index funds track. We then match index fund fees with our active funds based on their respective benchmarks.[17] The difference between TER and index fee is often called excess fee. Due to dispersion in the frequency of portfolio reporting intra-year, we use the average active share for fund $i$ in year $t$ in the denominator.

### 3.3 Empirical strategy

To examine whether investors can identify and respond to factors affecting a fund's ex-ante potential to outperform its benchmark index, we investigate the interaction between active fee (AF), excess fee (EF) and active share (AS) with future flows. Our main focus is on active fee. Therefore, if investors identify and respond to changes in active fee, then $\text{AF}_{i,t-1}$ and $\text{NF}_{i,t}$ should be negatively related. Since we want to measure the investors' response (net flow) to a *change* in either of the three variables discussed, we specify the following baseline regression:[18]

[16] Active share is

$$\text{AS}_{i,t} = \frac{1}{2}\sum_{j=1}^{N_{i,t}} |w_{j,i,t} - w_{j,b,t}|$$

where $w_{j,i}$ is the weight of security $j$ in fund $i$'s portfolio at time $t$ and $w_{j,b,t}$ is security $j$'s weight in the benchmark $b$ at time $t$. The weight LS fund's weight is equal to the overall fund's active share and the index fund's weight equals 1 - active share.

[17] If a benchmark-year fee is missing, we substitute in the average domicile-year index fee. If both are missing, we use the global-year index fee.

[18] We label the slope coefficient capturing the active fee-flow relation $\gamma_{AF}$, the slope coefficient for excess fee-flow $\gamma_{EF}$ and active share-flow $\gamma_{AS}$.

$$\mathrm{NF}_{i,t} = \varphi_{k,i} + \gamma_k \mathrm{X}_{k,i,t-1} + \theta \mathbf{Z}_{i,t-1} + \varepsilon_{k,i,t} \tag{9}$$

where $\mathrm{NF}_{i,t}$ is fund $i$'s net flow at time $t$. Furthermore, since we have three independent variables of interest, we denote that $\mathrm{X}_{k,i,t-1}$ can be main variable $k$ for fund $i$ at time $t-1$, where $k \in \{AF, EF, AS\}$. Since we model active capital allocation as a function of time-varying active fee, our primary interest lies in the time-series relationship between the two. Therefore, Equation 9 includes the $i$ subscript on the constant term $\varphi_{k,i}$, indicating that we do not require equal intercepts across funds. The main result of this paper is thus the estimated relationship between active fee ($k = AF$) and net flow from a fund fixed effects panel regression.[19] As fund fixed effects isolate the time variation, using time fixed effects alone is equivalent to using only cross-sectional variation in active fee and net flow.[20]

A restriction we do impose is that $\gamma_{k,1} = \gamma_{k,2} = \cdots = \gamma_k$, which increase the power of our inference. For an individual fund, a negative $\gamma_{\mathrm{AF},i}$ reflects investors' reallocation of capital to (from) fund $i$ in period $t$ based on their ability to identify a decrease (increase) in that fund's active fee (better upside potential) in period $t-1$. With a maximum of 13 yearly observations for each fund, the individual $\hat{\gamma}_{k,i}$'s would have low precision. Consequently, our focus is on $\hat{\gamma}_k$.[21]

Moreover, detecting this active fee-flow relationship demands variation in the upside potential of each fund. Thus, if a fund does not change its fee or varies its activity level when we observe them, they will not participate in our results. Therefore, a potential concern is that our results depend on only a small sub-sample of funds. Still, we detect only eight with zero variation in active fee, alleviating our worries.

$\mathbf{Z}_{i,t-1}$ is a vector of control variables, see Table A.1 for their definitions. In addition to those, different levels of fixed effects (e.g., domicile and benchmark) can be contained in $\mathbf{Z}_{i,t-1}$ as a way of controlling for unobserved group heterogeneity, as advocated by Gormley and Matsa (2014) amongst others. In all tables, we specify the level of variation used to estimate the reported coefficients.

[19] Using fund fixed effects isolate the time-series variation and excludes the cross-sectional variation when using OLS to estimate $\gamma$, as shown by Pástor, Stambaugh and Taylor (2017). They derive how $\hat{\gamma}$ is a weighted average of individual fund time-series regressions. Each fund's weight is influenced by the number of observations and the variation in active fee. The OLS estimate is given by

$$\hat{\gamma}_k = \sum_{i=1}^{N} \omega_i \hat{\gamma}_{k,i} \text{ where } \omega_{k,i} = \frac{T_i \sigma^2_{\mathrm{X}_{k,i}}}{\sum_{n=1}^{N} T_n \sigma^2_{\mathrm{X}_n}}.$$

With yearly observations on active fee and net flow, the number of observations for each fund falls within a relatively tight range. Consequently, the variability in active fee will be the key factor in each fund's contribution to the overall estimate.

[20] Here, the $\hat{\gamma}$ coefficient becomes a weighted average of yearly cross-sectional OLS regressions (see Pástor et al. (2017) for details).

[21] We do however note that in fund-by-fund regressions of Equation 9 where $\mathbf{Z}_{\mathrm{AF},i,t-1} = \varnothing$, we find that 58 % of the OLS slopes estimates $\hat{\gamma}_{\mathrm{AF},i}$ are negative. Furthermore, 7.3 % (2.4 %) are significantly negative at the 5 % (1 %) confidence level.

### 3.4 Descriptive statistics

In this section, we present descriptive statistics for our main variables across the entire sample and by country. For the control variables that are in $\mathbf{Z}_{i,t-1}$, we present descriptive statistics across the sample.

**Table 1:**
**Country descriptive statistics**

This table presents summary statistics on the number of actively managed equity mutual funds, the total net assets as of December 2019, the average net flow and average active and excess fee, as well as average active share by country for our sample.

| Country | Number of funds | N fund-year | TNA (USDbn) | Net flow (%) | Active fee (%) | Excess fee (%) | Active share (%) |
|---|---|---|---|---|---|---|---|
| Australia | 146 | 1,503 | 21 | 3.5 | 0.8 | 0.5 | 69 |
| Austria | 231 | 2,387 | 13 | -2 | 1.6 | 1.1 | 75 |
| Belgium | 237 | 2,098 | 6.2 | -5.7 | 1.3 | 0.8 | 68 |
| Canada | 1,399 | 11,946 | 269 | 3.8 | 2 | 1.6 | 83 |
| China | 25 | 287 | 6.9 | -0.9 | 1.2 | 1 | 86 |
| Denmark | 199 | 2,139 | 30 | -3.7 | 1.4 | 0.8 | 69 |
| Finland | 195 | 2,000 | 24 | 0.4 | 1.6 | 1.1 | 77 |
| France | 948 | 9,415 | 84 | -0.3 | 1.5 | 1.1 | 75 |
| Germany | 493 | 4,328 | 118 | -4 | 1.4 | 0.9 | 75 |
| Hong Kong | 134 | 1,346 | 35 | -1.3 | 1.2 | 0.9 | 74 |
| India | 319 | 3,268 | 91 | 3.6 | 2.2 | 1.5 | 74 |
| Ireland | 448 | 3,720 | 62 | 4.3 | 1.5 | 1.1 | 77 |
| Italy | 174 | 1,306 | 9.3 | -8.5 | 2.2 | 1.3 | 69 |
| Japan | 1,377 | 11,783 | 90 | -13 | 1.3 | 0.9 | 77 |
| Liechtenstein | 150 | 1,353 | 4.4 | -0.1 | 1.8 | 1.5 | 85 |
| Luxembourg | 76 | 855 | 25 | 1.3 | 1.2 | 0.9 | 79 |
| Malaysia | 244 | 2,460 | 7.4 | -2 | 0.7 | 0.5 | 83 |
| Netherlands | 95 | 1,006 | 18 | -6.1 | 0.8 | 0.5 | 74 |
| Norway | 153 | 1,632 | 36 | 1.3 | 1.3 | 0.9 | 78 |
| Portugal | 26 | 319 | 0.5 | -8.3 | 1.6 | 1 | 72 |
| Singapore | 143 | 1,400 | 9.4 | -8.4 | 1.4 | 1.1 | 80 |
| South Africa | 252 | 2,398 | 24 | 3.7 | 1.5 | 1 | 73 |
| South Korea | 924 | 7,148 | 20 | 0 | 1.3 | 1 | 78 |
| Spain | 416 | 3,357 | 28 | -2.5 | 1.8 | 1.1 | 70 |
| Sweden | 198 | 1,814 | 70 | 2.2 | 1.3 | 0.8 | 66 |
| Switzerland | 279 | 2,723 | 21 | -4.3 | 1.6 | 1 | 68 |
| Taiwan | 363 | 3,452 | 14 | -15 | 2.1 | 1.7 | 82 |
| Thailand | 320 | 3,000 | 23 | 0.8 | 1.3 | 0.8 | 72 |
| UK | 1,357 | 12,308 | 393 | -0.3 | 1.2 | 0.9 | 78 |
| USA | 2,010 | 19,421 | 1,564 | -1.7 | 0.5 | 0.4 | 82 |
| **All countries** | **13,331** | **122,172** | **3,119** | **−2.1** | **1.4** | **1** | **76** |

Table 1 illustrates the large variation in both active fee and net flows across our global sample of mutual funds. At one end of the range, we have countries such as Canada and Ireland that have experienced an average net inflow of capital over the sample period, while countries such as Japan and Taiwan have large average outflows. Across countries, we see a negative trend in an average country net outflow of 2.1 % per year over the sample period. This decline is similar to other sources; see, for example, Figure 3.14 in Factbook (2021). The Active fee is, as net flow, widely dispersed, ranging from an average of 0.5 % in the USA to 2.2 % in Italy, i.e., a spread of 170 bps between the countries. Overall, the average level of active fee is 1.4 % per year. For excess fee, the total cost of active management minus indexing, the global country average is 1 % per year. Also, here, the USA is the country with the lowest fee, charging on average only 40 bps per year for active management. Active share at the country level is in the interval between 66 % (Sweden) and 86 % (China).

In Table 2, we summarise our main variables as well as our control variables across all funds

**Table 2:**
**Variables descriptive statistics**

This table presents summary statistics for our main variables and additional control variables used in panel regressions. For each variable, we report the number of fund-year observations (N), mean and medians, standard deviation (SD) and mean average deviations (MAD), as well as the minimum and maximum values. These values are derived from the full set of observations in our sample. The variables are described in Table A.1 in Appendix A.1

| Variable | N | Mean | Median | SD | MAD | Min | Max |
|---|---|---|---|---|---|---|---|
| Net flow (%) | 122,172 | -1.99 | -4.94 | 47 | 19 | -100 | 1,182 |
| Active fee (%) | 122,172 | 1.34 | 1.27 | 0.99 | 0.83 | -2.9 | 9.39 |
| Excess fee (%) | 122,172 | 0.98 | 0.97 | 0.69 | 0.6 | -1.42 | 7.11 |
| Active share (%) | 122,172 | 78 | 81 | 19 | 20 | 3.94 | 100 |
| Gross alpha (%) | 122,172 | -0.83 | -0.81 | 7.94 | 5.48 | -48 | 85 |
| TNA (USDm) | 122,172 | 484 | 65 | 3,215 | 85 | 1 | 185,786 |
| Family TNA (USDbn) | 122,172 | 16 | 2.29 | 70 | 3.18 | 0 | 1,976 |
| Fund age | 122,172 | 12 | 10 | 9.92 | 8.15 | 0.11 | 94 |
| Industry size (USDbn) | 122,172 | 787 | 136 | 1,518 | 167 | 1.63 | 6,712 |
| GDP per capita (USDk) | 122,172 | 43 | 44 | 21 | 11 | 0.81 | 181 |
| Competition (1-HHI) | 122,172 | 94 | 96 | 7.52 | 2.62 | 8.92 | 100 |
| Years of schooling | 122,172 | 12 | 12 | 1.69 | 1.19 | 5 | 14 |
| Market share index funds (% of TNA) | 122,172 | 17 | 14 | 14 | 11 | 0.02 | 83 |

and years. We find that the average fund has a negative yearly net flow at -1.99%, charges 1.34% for active management, is 12 years old with 484 USDm in assets and delivers a negative gross alpha of -83 bps per year.

# 4 Main results

In this section, we report the estimated slope coefficients on active fee ($\hat{\gamma}_{AF}$), excess fee ($\hat{\gamma}_{EF}$) and active share ($\hat{\gamma}_{AS}$) for various specifications of the panel regression capturing the ex-ante upside potential-flow relation. All tests, follow the structure from Equation 9.

## 4.1 Active fee and flow

Our first hypothesis relates to the central implication from the model in Section 2 of a negative time-series relation between active fee and subsequent flows. In Table 3, we provide details on this relation by reporting the estimated slope coefficient of active fee, or $\hat{\gamma}_{AF}$, for variations of Equation 9. In the first four columns, control variables are not included, while they are included in Columns 5-8.

Column 1 reports the slope coefficient when no fixed effects are included in the panel regression. This entails that we have dropped the $i$ subscript from the constant term in Equation 9, thus we are not only imposing the restriction that $\gamma_{AF,1} = \gamma_{AF,2} = \cdots = \gamma_{AF}$, but also that $\varphi_{AF,1} = \varphi_{AF,2} = \cdots = \varphi_{AF}$. The estimate in Column 1 thus reflects both cross-sectional and time-series variation. The estimate, -0.76 is statistically significant. Column 2 is the result from a purely cross-sectional regression, in which fund fixed effects in Equation 9 $\varphi_{AF,i}$ is replaced by year fixed effects $\mu_{AF,t}$. The estimate of -1.40 means that isolating cross-sectional variation strengthens

**Table 3:**
**Active fee-flow relation**

This table reports the estimated slope coefficients from eight different panel regressions of $NF_{i,t}$ on $AF_{i,t-1}$ They differ in their treatment of fixed effects and inclusion of additional controls beyond active fee. The control variables are described in Table A.1 in Appendix A.1. $NF_{i,t}$ is fund $i$'s net flow in year $t$. $AF_{i,t-1}$is fund $i$'s active fee in the year ending before net flow is measured. For each model we report the level of fixed effects and the dimension we cluster our standard errors. As short-hand notation, we use F for fund and Y for year. Asterisks denote statistical significance: $^{***}$p<0.01, $^{**}$p<0.05, $^{*}$p<0.1.

| | Net flow (t) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Active fee$_{t-1}$ | $-0.76^{***}$ | $-1.40^{***}$ | $-2.98^{***}$ | $-3.03^{***}$ | $-0.13$ | $-1.36^{***}$ | $-2.81^{***}$ | $-4.21^{***}$ |
| | (0.13) | (0.33) | (0.58) | (0.60) | (0.14) | (0.33) | (0.55) | (0.56) |
| Controls | No | No | No | No | Yes | Yes | Yes | Yes |
| Fixed effects | None | Y | F | F + Y | None | Y | F | F + Y |
| SE | Robust | F + Y | F + Y | F + Y | Robust | F + Y | F + Y | F + Y |
| N | 94,571 | 94,571 | 94,571 | 94,571 | 86,186 | 86,186 | 86,186 | 86,186 |
| $R^2$ | 0.00 | 0.00 | 0.25 | 0.25 | 0.03 | 0.03 | 0.34 | 0.34 |
| Adjusted $R^2$ | 0.00 | 0.00 | 0.13 | 0.13 | 0.03 | 0.03 | 0.23 | 0.23 |

the active fee-flow relation considerably.

The slope coefficient in Column 3 is negative and highly significant. This finding of a negative active fee-net flow relation in the time-series is the main empirical result of the paper. The average within-fund standard deviation of $X_{AF,i,t-1}$ is 0.28. Therefore, the estimated slope of -2.98 implies that a one standard deviation increase in a fund's active fee translates to an decrease in annual net flow of 83.4 bps ($-0.834\% = -2.98 \cdot 0.28$). This number is economically significant in that it involves a 42 % (= 0.834/1.99) reduction in unconditional expected yearly outflow.

In Column 4, we add time fixed effects to the panel regression, meaning that the matrix of controls $\mathbf{Z}_{i,t-1}$ is replaced by a vector of year-dummies $\mu_t$. The resulting slope coefficient is slightly more negative, -3.03. The only thing separating the estimate in Columns 3 and 4 is the inclusion of year fixed effects. This addition controls for any unobserved variables that change over time but not across funds within each sample country.[22] This can be macroeconomic variables, regulatory changes etc. Because our results with and without time fixed effects are close to each other, such aggregate variables cannot explain the negative relation between active fee and net flow.

In Column 5-8, we run the same panel regressions as in Column 1-4, but now we condition on the set of control variables listed in Table A.1. Most importantly, the main specification, fund fixed effects, is only somewhat weakened when controlling for fund and country characteristics. The economics of this slope coefficient suggests that a one standard deviation increase in active

[22]In single country studies, adding time fixed effects, control for unobserved variables that change over time but not across funds within that country. Thus, time fixed effects can be interpreted as country-specific unobserved time-varying changes affecting all funds equally. In our case, adding time fixed effects, in the same way, controls for global unobserved time-varying changes that affect all funds in the same way. Therefore, when including time fixed effects, we first residualize the main variable of interest by regressing it on year dummies country by country. This variable stripped of time-varying unobserved effects is then used in conjunction with fund fixed effects in the panel regressions.

fee is associated with a subsequent decrease in the net flow of 78.7 bps ($-0.787\% = -2.81 \cdot 0.28$). This effect is still economically important as it represents a 39.5% reduction in unconditional expected yearly outflow. Using both cross-sectional and time-series variation in combination with controls weakens the relation a lot. In Column 6, we see that the pure cross-sectional relationship still obtains. More interestingly, controlling for country-specific unobserved time-varying effects by adding year fixed effects in Column 8 increases the estimated slope coefficient in absolute terms. Using the within-fund standard deviation of active fee after controlling for year fixed effects, we see it decrease marginally from 0.28 to 0.275. This entails that a one standard deviation increase in active fee is related to a 116 bps ($-1.16\% = -4.21 \cdot 0.275$) decrease in the year-ahead net flows, a reduction in unconditional expected outflow of 58.3% (= 1.16/1.99). Thus, consistent with the hypothesis, we find that a fund's active fee negatively predicts the fund's subsequent flow.

### 4.2 Excess fee and flow

Having established a strong time-series relation between active fee and net flow in Section 4.1, we now zoom in on the numerator of the active fee measure, excess fee. Our hypothesis is based on Equation 3, where we show how that higher fees lead to lower value-added. As such, we test whether there is a negative relation between excess fee and net flow. Following the same setup as before, we now run our panel regressions on excess fee instead.

**Table 4:**
**Excess fee-flow relation**

This table reports the estimated slope coefficients from eight different panel regressions of $\mathrm{NF}_{i,t}$ on $\mathrm{EF}_{i,t-1}$. They differ in their treatment of fixed effects and inclusion of additional controls beyond excess fee. The control variables are described in Table A.1 in Appendix A.1. $\mathrm{NF}_{i,t}$ is fund $i$'s net flow in year $t$. $\mathrm{EF}_{i,t-1}$ is fund $i$'s excess fee in the year ending before net flow is measured. For each model we report the level of fixed effects and the dimension we cluster our standard errors. As short-hand notation, we use F for fund and Y for year. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1.

| | Net flow (t) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Excess fee$_{t-1}$ | −0.18 | −1.20** | −1.93* | −0.34 | 0.11 | −2.22*** | −2.09** | −5.26*** |
| | (0.18) | (0.47) | (1.07) | (1.00) | (0.22) | (0.50) | (0.96) | (0.97) |
| Controls | No | No | No | No | Yes | Yes | Yes | Yes |
| Fixed effects | None | Y | F | F + Y | None | Y | F | F + Y |
| SE | Robust | F + Y | F + Y | F + Y | Robust | F + Y | F + Y | F + Y |
| N | 94,571 | 94,571 | 94,571 | 94,571 | 86,186 | 86,186 | 86,186 | 86,186 |
| $R^2$ | 0.00 | 0.00 | 0.25 | 0.25 | 0.03 | 0.03 | 0.34 | 0.34 |
| Adjusted $R^2$ | 0.00 | 0.00 | 0.13 | 0.13 | 0.03 | 0.03 | 0.23 | 0.23 |

Table 4 list our estimated slope coefficients for the various specifications. The results vary to a larger extent, but are for the most part, in agreement with what the literature has found earlier (Sirri and Tufano (1998), and Barber et al. (2005)). There is no significant relation between excess fee when employing cross-sectional and time-series variation (Column 1), nor when controlling for year fixed effects in the time-series relation (Column 4). The pure time-series relation is statistically significant on the 10% level, and the relation is here weaker than the comparable one in Table 3. Despite being on the statistically weaker end, the economic significance is not trivial. Given a within-fund standard deviation in excess fee of 0.185, the coefficient in Column 3 suggests that one standard deviation change is associated with a subsequent change flow of 35.7 bps, which translates to an 18% relative change compared to the unconditional average.

Including controls in Column 5-8 brings the slope coefficients more in line with the results in the previous section. The pure cross-section results are stronger for excess fee ($\hat{\gamma}_{EF} = -2.22$ than for active fee ($\hat{\gamma}_{AF} = -1.36$), a little weaker for in the time-series ($\hat{\gamma}_{EF} = -2.09$ vs. $\hat{\gamma}_{AF} = -2.81$) and considerably stronger when controlling for year fixed effects in the time-series relationship. In sum, we find a negative effect from excess fee, however, not as strong as suggested by the model.

### 4.3 Active share and flow

Lastly, in this section, we test the relation between active share and net flow. We expect $\hat{\gamma}_{AS}$ to be positive, indicating that an increase in the size of the long-short fund component attracts capital. As above, we test this relation using variants of Equation 9.

**Table 5:**
**Active share-flow relation**

This table reports the estimated slope coefficients from eight different panel regressions of $\text{NF}_{i,t}$ on $\text{AS}_{i,t-1}$. They differ in their treatment of fixed effects and inclusion of additional controls beyond active share. The control variables are described in Table A.1 in Appendix A.1.$\text{NF}_{i,t}$ is fund $i$'s net flow in year $t$. $\text{AS}_{i,t-1}$ is fund $i$'s average active share in the year ending before net flow is measured. For each model we report the level of fixed effects and the dimension we cluster our standard errors. As short-hand notation, we use F for fund and Y for year. Asterisks denote statistical significance: $^{***}$p<0.01, $^{**}$p<0.05, $^{*}$p<0.1.

| | Net flow (t) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Active share$_{t-1}$ | 0.07$^{***}$ | 0.06$^{***}$ | 0.21$^{***}$ | 0.30$^{***}$ | 0.01 | −0.00 | 0.16$^{***}$ | 0.15$^{***}$ |
| | (0.01) | (0.01) | (0.03) | (0.03) | (0.01) | (0.02) | (0.03) | (0.02) |
| Controls | No | No | No | No | Yes | Yes | Yes | Yes |
| Fixed effects | None | Y | F | F + Y | None | Y | F | F + Y |
| SE | Robust | F + Y | F + Y | F + Y | Robust | F + Y | F + Y | F + Y |
| N | 94,571 | 94,571 | 94,571 | 94,571 | 86,186 | 86,186 | 86,186 | 86,186 |
| $R^2$ | 0.00 | 0.00 | 0.26 | 0.26 | 0.03 | 0.03 | 0.34 | 0.34 |
| Adjusted $R^2$ | 0.00 | 0.00 | 0.13 | 0.13 | 0.03 | 0.03 | 0.23 | 0.23 |

In Table 5, we see that the sign of $\hat{\gamma}_{AS}$ matches our ex-ante expectations. Furthermore, no matter what source of variation we draw from, the relation is highly significant when pooling across funds. Cremers et al. (2016) look at a similar relation but rely on the time-series variation at the benchmark level. In contrast, we use time-series variation at the fund level, meaning that our coefficient estimates are not directly comparable.

Column 1-4 also shows the time-series relation is more than 3x larger in magnitude than the cross-sectional association. The slope estimate in Column 3 suggest that a one (within-fund) standard deviation increase in active share increases expected net flow by 110 bps ($1.1\% = 0.21 \cdot 5.25$). This effect is on par with the effect found in Column 8 in Section 4.1, where we controlled for both year fixed effects and other fund and country-level factors. The slope estimate increases from Column 3 to 4, where we also control for time-varying unobserved variables, but not by much. Therefore, we infer that these types of unobserved variables are not what is driving our results.

Conditioning the expected net flow on additional control variables weakens the active share -flow relation in the interaction between the cross-section and time-series in Column 5, as well as the purely cross-sectional relation in Column 6, but it persists for our main configuration in Column 7, even though the economic impact is reduced.[23] In contrast to the large change in slopes when transitioning from Column 7 to 8 in Table 4, controlling for time fixed effects in the active share-flow relation hardly changes the estimate, meaning that these unobserved aggregate variables cannot explain the positive relation between active share and net flow.

To assess the statistical significance of the upside potential-net flow slope estimates, we cluster independently by fund and country·year. Our rationale for clustering by fund is that net flows are shown to be persistent by Coval and Stafford (2007). Furthermore, we cluster by domicile·year to allow for cross-sectional dependence in net flows within each country.

We have established a clear negative time-series relationship between ex-ante upside potential for outperformance and ensuing net flows. The relationship obtains independently of how we measure this potential, i.e., by active fee, excess fee or active share. Even though we focus on the time-series relation, we see in Appendix B that the same result also hold in the cross-section. Furthermore, the time-series relationship is consistent whether we pool across all funds or by conditioning on additional control variables. Controlling for aggregate and unobserved time-varying factors alter, for the most part, not the relationship substantially. Where it does change, it only strengthens our results, and these changes are driven by the fact that some aggregate time-varying factors affect the level of excess fee in different countries. The documented relations are not just highly significant in a statistical fashion, but also economically meaningful in that conditioning on them reduces next year's expected net outflow by 35.7% (excess fee) and 1.1% (active share) when pooled across all funds, and a reduction between 19.6% (excess fee) and 42.2% (active share) as compared to the unconditional expected net flow.

# 5 Morningstar ratings and the active fee-fund flow relation

This section extends the main results for the active fee-flow relation by testing how Morningstar ratings influence the relation between active fee and fund flow. The Morningstar Rating, or also known as the "star rating," is a purely quantitative, backward-looking measure of a fund's past

[23]A one standard deviation increase in active share is now only associated with an increase in expected net flow of 84 bps, which is on par with the result for active fee, Column 3 Table 3 in Section 4.1.

performance, measured from one to five stars, and are calculated each month.[24] Ratings are annualized by calculating the average monthly rating within a year for each fund. For a more detailed explanation of the Morningstar ratings and methodology, see Morningstar Ratings.

Our tests consist of three different regressions specifications. First, we test how the Morningstar ratings are related to active fee by running the regression in Equation (9) with ratings as the dependent variable. This specification assesses whether active fee is orthogonal to the Morningstar ratings or whether they measure the same thing.

Table 6 report the results for all regression specifications. In Column 1-2, we report the estimated slope coefficient on active fee, or $\hat{\gamma}_{AF}$, on Morningstar ratings. We find a small and negative relationship between these two variables. However, the economic significance of the results is low. A one standard deviation increase in active fee is associated with a decrease of $-0.006$ $(-0.27 \cdot 0.024)$ stars in the rating, where the rating ranges from 1 to 5. We also note that the coefficients do not change much when adding the domicile-year fixed effects in Column 2 and that the standard errors are fairly stable from Column 1 to Column 2 when adding double cluster. These results show that active fee and ratings are mainly independent of each other.

**Table 6:**
**Active fee-flow relation and Morningstar ratings**

This table reports the estimated slope coefficients for active fee from six different panel regressions. In Columns 1 and 2, we estimate the relationship between Morningstar ratings and active fee. In Columns 3 and 4, we estimate the active fee-flow relationship when controlling for Morningstar ratings. Columns 5 and 6 estimate differences in active fee-flow relationships for funds with high and low ratings. Fixed effects and level of cluster robust standard errors are denoted at the bottom of the table. Asterisks denote statistical significance: $^{***}$p<0.01, $^{**}$p<0.05, $^{*}$p<0.1

| | Morningstar Rating (t) | | Net flow (t) | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| $AF_{t-1}$ | $-0.02$ | $-0.02^{*}$ | $-2.99^{***}$ | $-3.39^{***}$ | $-3.29^{***}$ | $-3.76^{***}$ |
| | (0.01) | (0.01) | (0.47) | (0.57) | (0.50) | (0.60) |
| $Rating_{t-1}$ | | | $6.19^{***}$ | $6.38^{***}$ | | |
| | | | (0.32) | (0.43) | | |
| $High_{t-1}$ | | | | | $4.97^{***}$ | $5.04^{***}$ |
| | | | | | (0.81) | (0.95) |
| $AF_{t-1} \cdot High_{t-1}$ | | | | | 0.20 | 0.39 |
| | | | | | (0.61) | (0.63) |
| $Low_{t-1}$ | | | | | $-6.78^{***}$ | $-6.52^{***}$ |
| | | | | | (0.88) | (0.92) |
| $AF_{t-1} \cdot Low_{t-1}$ | | | | | $1.52^{***}$ | $1.45^{***}$ |
| | | | | | (0.46) | (0.49) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Fixed effects | F | F + Y | F | F + Y | F | F + Y |
| SE | F | F + Y | F | F + Y | F | F + Y |
| N | 53,593 | 53,593 | 53,974 | 53,974 | 53,974 | 53,974 |
| $R^2$ | 0.71 | 0.72 | 0.33 | 0.36 | 0.32 | 0.35 |
| Adjusted $R^2$ | 0.65 | 0.66 | 0.20 | 0.23 | 0.19 | 0.22 |

Second, we add lagged Morningstar ratings as a control variable to the main regression in Equation (9). We argue that by controlling for the ratings, we improve the identification in two

[24]For more information about the Morningstar Star Ratings see: Morningstar Ratings.

ways. First, it extracts the effect of ratings, giving a more clean estimate of the active fee-flow relationship. Second, it allows for interpreting the economic significance of the active fee-flow relationship relative to the rating-flow relationship.

Column 3-4 report the estimated slope coefficient of active fee, or $\hat{\gamma}_{AF}$, on net flows in a model where we control for the lagged Morningstar ratings. In the model with fund fixed effects (Column 3), the active fee-flow relation also holds when controlling for the ratings, with significance at the 1% level. Compared to our main results in Column 7 in Table 3, the coefficients do not change much. A one standard deviation increase in active fee translates into an decrease in net flow of 0.81% ($-2.990 \cdot 0.27$). When adding year fixed effects and two-way clusters (Column 4), our findings are similar and still significant at the 1% level. Compared to Column 8 in Table 3, the magnitude of the coefficient drops slightly. A one standard deviation increase in active fee translates into a decrease in net flow of 0.92% ( $-3.394 \cdot -0.27$). Since past returns are part of control variables and correlated with ratings, we also perform estimation without these control variables. However, this specification does not change the estimated results.

In terms of economic significance, we argue that our findings from Section 4 hold and that active fee still explains a fairly large portion of the flows. From Columns 3 and 4, we find that a one standard deviation increase in Morningstar rating leads to a rise of 3.28% ($0.53 \cdot 6.195$) and 3.38% ($0.53 \cdot 6.378$) in net flow. This implies that the Morningstar ratings are more important for explaining the flows, but we argue that active fee still explains a fairly large portion of the flows. These results also confirm the findings from the literature showing that Morningstar ratings are important for investors when choosing funds. However, we claim that the active fee signal is more rational than the Morningstar ratings.

Our third set of tests are related to high and low ratings for funds. To examine differences in the active fee-flow relation across funds based on ratings, we add interactions of high-rated and low-rated funds with active fee to the regression in Equation (9). We define high-rated funds as funds with 4 or 5 stars and low-rated funds as funds with 1 or 2 stars. This implies that the base of the regression is medium-rated funds with 3 stars.

Columns 5 and 6 show that the relation between active fee and fund flows holds for the 3-star funds. The active fee-flow relationships are less negative for low-rated funds compared to the middle group. For high-rated funds, the interaction term is not statistically different from the funds with medium rating. In the hypothesis development part, we suggest that the less sensitive to changes in active fee for low-rated funds might be caused by less sophisticated investors for these funds. This characteristic implies that they are less inclined to collect information and respond to changes in active fee.

Based on our findings in this section, we argue that active fee as a flow signal is independent of Morningstar ratings. The implies that even though Morningstar ratings are well-grounded in the industry and easily accessible for investors, the signal from active fee is still followed by investors. Relative to our findings in Section 4.1, the magnitude of the coefficient of the active fee-flow relation does not change much, and the statistical significance holds. We also find that investors in low-rated funds are less likely to use active fee to allocate assets, which we connect to low financial sophistication.

## 6 Fund characteristics and the active fee-fund flow relation

In this section, we test our third set of hypotheses. We explore whether different fund characteristics influence the active fee-flow relation. We start with the influence of past performance before we explore who fund size interfere before we finish with fund styles. Table 7 report the results. We note from the first row in all regressions that the negative and statistically significant active fee-flow relation holds at the 1% level for the middle groups for all of the characteristics.

In general, fund performance is not very persistent (see Carhart (1997)). Despite this lack of persistence, multiple studies document that fund flows follows the funds with highest past performance (see e.g., Ippolito (1992), Chevalier and Ellison (1997), Sirri and Tufano (1998), and Del Guercio and Tkac (2002)). Ferreira et al. (2012) show in an international study that the flow-performance relationship is most present in developed countries. Based on this, we first group the funds based on past performance. As Morningstar ratings are based on past performance, these tests are highly related to the tests in the previous section, but past returns are not as accessible as the Morningstar ratings. We group the funds into terciles based on the benchmark adjusted net returns of the previous year.[25] Funds are sorted within each country-benchmark segment to determine the rankings.

In Columns 1-2, we show the result where funds are ranked based on past performance. Similar to Morningstar ratings, we find that the active fee-flow relation is less negative for the worst-performing funds (lowest tercile).[26] In the group of top performers, the coefficient of the interaction variable is not statistically significant, but the relationship is negative. Since Morningstar ratings are based on past performance, the similarities between these tests do not come as a surprise.

We also note that the coefficients of the stand-alone dummy variables for $High_{t-1}$ and $Low_{t-1}$, which can be interpreted as the average flow into these funds relative to the middle group of funds, are highly significant and of opposite signs. These findings confirm the flow-performance relationship of previous studies, where we find that money flows into the top performers and out of the bottom performers.

Next, we investigate the influence of fund size. As suggested by the theoretical model of Berk and Green (2004), funds have decreasing returns to scale. This is also documented in several empirical studies in the mutual fund literature (see e.g., Pástor, Stambaugh and Taylor (2015) and Zhu (2018)). We can relate this feature to our setting, with a lower skill for large funds than small funds. This relation implies a less negative relation between active fee and fund flow for larger funds. For an investor, it is more important that the fund is active when the manager is skilled than not, i.e., more important for a small fund than a large fund. To test this prediction,

[25] We use benchmark adjusted net returns as these are the returns left for investors.

[26] The number of observations with Morningstar data is lower due to missing data or that we couldn't match the whole sample.

## Table 7:
## Active fee-flow relation across funds

This table reports the estimated slope coefficients of active fee from eight different panel regressions. Columns 1 and 2 estimate differences in active fee-flow relationships for funds with high and low past performance. Columns 3 and 4 estimate differences in active fee-flow relationships for funds small and large funds. Column 5 and 6 estimate differences in active fee-flow relationships between small-cap and large-cap funds. Columns 7 and 8 estimate differences in active fee-flow relationships for value and growth funds. Fixed effects and level of cluster robust standard errors are denoted at the bottom of the table. Asterisks denote statistical significance: ***p<0.01, **p<0.05, *p<0.1

| Dependent variable | Net flow (t) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Characteristic | High-low performance | | Large-small funds | | Large-cap - small-cap | | Value-growth | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| $AF_{t-1}$ | −3.24*** | −4.26*** | −2.33*** | −3.52*** | −2.62*** | −2.69*** | −3.37*** | −3.87*** |
| | (0.46) | (0.57) | (0.51) | (0.61) | (0.88) | (1.02) | (0.67) | (0.73) |
| $Low_{t-1}$ | −4.71*** | −4.64*** | | | | | | |
| | (0.55) | (0.67) | | | | | | |
| $High_{t-1}$ | 3.52*** | 3.60*** | | | | | | |
| | (0.55) | (0.70) | | | | | | |
| $AF_{t-1} \cdot Low_{t-1}$ | 1.13*** | 1.09*** | | | | | | |
| | (0.30) | (0.34) | | | | | | |
| $AF_{t-1} \cdot High_{t-1}$ | −0.06 | −0.17 | | | | | | |
| | (0.34) | (0.39) | | | | | | |
| $Small_{t-1}$ | | | 19.72*** | 19.66*** | | | | |
| | | | (1.12) | (1.21) | | | | |
| $Large_{t-1}$ | | | −16.79*** | −17.20*** | | | | |
| | | | (0.90) | (1.41) | | | | |
| $AF_{t-1} \cdot Small_{t-1}$ | | | −0.99 | −0.81 | | | | |
| | | | (0.63) | (0.67) | | | | |
| $AF_{t-1} \cdot Large_{t-1}$ | | | 1.15** | 1.22** | | | | |
| | | | (0.51) | (0.56) | | | | |
| $\text{Small-cap}_{t-1}$ | | | | | | | | |
| $\text{Large-cap}_{t-1}$ | | | | | | | | |
| $AF_{t-1} \cdot \text{Small-cap}_{t-1}$ | | | | | 3.05 | 0.82 | | |
| | | | | | (1.87) | (1.95) | | |
| $AF_{t-1} \cdot \text{Large-cap}_{t-1}$ | | | | | −1.01 | −1.94 | | |
| | | | | | (1.02) | (1.21) | | |
| $AF_{t-1} \cdot Growth_{t-1}$ | | | | | | | 1.88* | 0.74 |
| | | | | | | | (0.99) | (1.08) |
| $AF_{t-1} \cdot Value_{t-1}$ | | | | | | | −1.19 | −1.10 |
| | | | | | | | (1.11) | (1.10) |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Fixed effects | F | F + Y | F | F + Y | F | F + Y | F | F + Y |
| SE | F | F + Y | F | F + Y | F | F + Y | F | F + Y |
| N | 82,827 | 82,827 | 82,827 | 82,827 | 66,534 | 66,534 | 66,534 | 66,534 |
| $R^2$ | 0.35 | 0.37 | 0.31 | 0.34 | 0.34 | 0.36 | 0.34 | 0.36 |
| Adjusted $R^2$ | 0.23 | 0.26 | 0.19 | 0.22 | 0.23 | 0.25 | 0.23 | 0.25 |

we sort funds into terciles based on the fund TNA at the end of the previous year within each domicile. Next, we run regressions with interactions between the bottom and top terciles, leaving the middle-sized funds as the regression base.

The difference in active fee-flow relation across funds based on size is presented in Columns 3 and 4. We find a positive and statistically significant coefficient of the interaction for large funds. This result shows, as predicted, that investors investing in large funds are less sensitive towards active fee. For small funds, we find a negative effect but not statistically significant difference from the mid-size funds. The results for these tests support our hypothesis. If skill is higher for smaller funds, investors should care more about active fee in these funds compared to larger funds.

Finally, we examine the Morningstar 3x3 style-boxes.[27] In the tests in Table 7, we focus on the styles independently, i.e. large vs. small-cap funds and value vs. growth funds. We are agnostic and do not form any specific predictions.

Columns 5-8 report the results for how fund styles influence the active fee-fund relation. For small vs. large-cap funds, we see no evidence of any difference in price sensitivity. We find a positive but weak statistically significant coefficient of the interaction between active fee and flow for growth funds. Regarding value funds, the coefficient is negative but statistically insignificant.

In sum, our findings in these tests show that there are differences in the active fee-fund flow relation dependent on different fund characteristics. In our model in Section 2 we show that rational investors base investment choices on active fee. We find that investors in the bottom tercile based on fund performance and in the top tercile based on fund size are less rational, i.e., follows active fee less. For the style characteristics from Morningstar, we find only weak evidence for differences across funds.

## 7 Conclusion

This paper develops a simple model showing that active fee can be a signal for future value creation and hypothesize that investors follow it when picking their mutual funds. Our empirical results are consistent with our predictions. For example, we show that a one standard deviation increase in active fee translates into a reduction of 83.4 basis points in the subsequent annual net flow. This result is both statistically and economically significant. Furthermore, the relation between active fee and flow comes from both the excess fee and the activity level. Thus, we show that in a market where it is hard for both researchers and investors to identify what creates value for the investors, at least some investors use a "rational" signal such as active fee.

However, we argue that active mutual fund investors should have responded more than they did. It is not easy to regulate financial markets, but maybe active fee should be part of the investor information. As suggested by Cremers and Curtis (2016), disclosure of active fee can help prevent closet-indexing.[28] Moreover, since Morningstar ratings are unrelated to active fee, a combination of both can be valuable. This question, we leave for future research.

[27]The nine styles in the style-box include small-cap value, small-cap blend, small-cap growth, mid-cap value, mid-cap blend, mid-cap growth, large-cap value, large-cap blend, and large-cap growth. See Morningstar Style-Box for more information about the Morningstar style-box.

[28]The alternative to disclosing this type of information ex-ante is a potentially costly ex-post intervention and correction by regulatory authorities. For an example of the latter, see Bjerksund, Doskeland, Sjuve and Ørpetveit (2020).

## References


Barber, B. M., Huang, X. and Odean, T. (2016), 'Which Factors Matter to Investors? Evidence from Mutual Fund Flows', *The Review of Financial Studies* **29**(10), 2600–2642.

Barber, B. M., Odean, T. and Zheng, L. (2005), 'Out of sight, out of mind: The effects of expenses on mutual fund flows', *The Journal of Business* **78**(6), 2095–2120.

Ben-David, I., Li, J., Rossi, A. and Song, Y. (2019), 'What do mutual fund investors really care about?', *Fisher College of Business Working Paper* **118**(2019-03), 005.

Berk, J. B. and Green, R. C. (2004), 'Mutual fund flows and performance in rational markets', *Journal of political economy* **112**(6), 1269–1295.

Berk, J. B. and Van Binsbergen, J. H. (2015), 'Measuring skill in the mutual fund industry', *Journal of Financial Economics* **118**(1), 1–20.

Berk, J. B. and van Binsbergen, J. H. (2016), 'Assessing asset pricing models using revealed preference', *Journal of Financial Economics* **119**(1), 1–23.

Betermier, S., Calvet, L. E. and Sodini, P. (2017), 'Who are the value and growth investors?', *The Journal of Finance* **72**(1), 5–46.

Bjerksund, P., Doskeland, T., Sjuve, A. W. and Ørpetveit, A. (2020), 'Forced to be active: Evidence from a regulation intervention', *Available at SSRN 3635718* .

Brown, D. C. and Davies, S. W. (2017), 'Moral hazard in active asset management', *Journal of Financial Economics* **125**(2), 311–325.

Brown, K. C., Tiu, C. and Yoeli, U. (2020), 'The decision to concentrate: Active management, manager skill, and portfolio size', *The Journal of Portfolio Management* **46**(5), 41–62.

Carhart, M. M. (1997), 'On persistence in mutual fund performance', *The Journal of finance* **52**(1), 57–82.

Carlin, B. I. (2009), 'Strategic price complexity in retail financial markets', *Journal of financial Economics* **91**(3), 278–287.

Chevalier, J. and Ellison, G. (1997), 'Risk taking by mutual funds as a response to incentives', *Journal of political economy* **105**(6), 1167–1200.

Cooper, M. J., Halling, M. and Yang, W. (2020), 'The persistence of fee dispersion among mutual funds', *Review of Finance* **25**(2), 365–402.

Coval, J. and Stafford, E. (2007), 'Asset fire sales (and purchases) in equity markets', *Journal of Financial Economics* **86**(2), 479–512.

Cremers, K. and Curtis, Q. (2016), 'Do mutual fund investors get what they pay for: Securities law and closet index funds', *Va. L. & Bus. Rev.* **11**, 31.

Cremers, K. M., Fulkerson, J. A. and Riley, T. B. (2019), 'Challenging the conventional wisdom on active management: A review of the past 20 years of academic literature on actively managed mutual funds', *Financial Analysts Journal* **75**(4), 8–35.

Cremers, K. M. and Petajisto, A. (2009), 'How active is your fund manager? a new measure that predicts performance', *The review of financial studies* **22**(9), 3329–3365.

Cremers, M., Ferreira, M. A., Matos, P. and Starks, L. (2016), 'Indexing and active fund management: International evidence', *Journal of Financial Economics* **120**(3), 539–560.

Del Guercio, D. and Tkac, P. A. (2002), 'The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds', *Journal of Financial and Quantitative Analysis* pp. 523–557.

Evans, R. B. and Sun, Y. (2021), 'Models or Stars: The Role of Asset Pricing Models and Heuristics in Investor Risk Adjustment', *The Review of Financial Studies* **34**(1), 67–107.

Factbook, M. F. (2021), 'Investment company institute'. Available at https://www.ici.org/system/files/2021-05/2021_factbook.pdf.

Ferreira, M. A., Keswani, A., Miguel, A. F. and Ramos, S. B. (2012), 'The flow-performance relationship around the world', *Journal of Banking & Finance* **36**(6), 1759–1780.

Gallaher, S., Kaniel, R. and Starks, L. T. (2006), 'Madison avenue meets wall street: Mutual fund families, competition and advertising', *Competition and Advertising (January 2006)* .

Gormley, T. A. and Matsa, D. A. (2014), 'Common errors: How to (and not to) control for unobserved heterogeneity', *The Review of Financial Studies* **27**(2), 617–661.

Grinold, R. C. and Kahn, R. N. (2000), *Active portfolio management*, McGraw Hill New York.

Huang, C., Li, F. and Weng, X. (2020), 'Star ratings and the incentives of mutual funds', *The Journal of Finance* **75**(3), 1715–1765.

Ippolito, R. A. (1992), 'Consumer reaction to measures of poor quality: Evidence from the mutual fund industry', *The Journal of Law and Economics* **35**(1), 45–70.

Markowitz, H. (1952), 'Portfolio selection', *The Journal of Finance* **7**(1), 77–91.

Marmoiton, A. (2021), 'An approximate formula linking active share and tracking error', *Available at SSRN 3771144* .

Merton, R. C. (1987), 'A simple model of capital market equilibrium with incomplete information', *The Journal of Finance* **42**(3), 483–510.

Müller, S. and Weber, M. (2010), 'Financial literacy and mutual fund investments: who buys actively managed funds?', *Schmalenbach Business Review* **62**(2), 126–153.

Nieuwerburgh, S. V. and Veldkamp, L. (2010), 'Information acquisition and under-diversification', *The Review of Economic Studies* **77**(2), 779–805.

Pástor, L., Stambaugh, R. F. and Taylor, L. A. (2015), 'Scale and skill in active management', *Journal of Financial Economics* **116**(1), 23–45.

Pástor, L., Stambaugh, R. F. and Taylor, L. A. (2017), 'Do funds make more when they trade more?', *The Journal of Finance* **72**(4), 1483–1528.

Pástor, L., Stambaugh, R. F. and Taylor, L. A. (2020), 'Fund tradeoffs', *Journal of Financial Economics* **138**(3), 614–634.

Roussanov, N., Ruan, H. and Wei, Y. (2021), 'Marketing mutual funds', *The Review of Financial Studies* **34**(6), 3045–3094.

Schumacher, D. (2018), 'Home bias abroad: Domestic industries and foreign portfolio choice', *The Review of Financial Studies* **31**(5), 1654–1706.

Sirri, E. R. and Tufano, P. (1998), 'Costly search and mutual fund flows', *The journal of finance* **53**(5), 1589–1622.

Zhu, M. (2018), ‘Informative fund size, managerial skill, and investor rationality’, *Journal of Financial Economics* **130**(1), 114–134.

Ľuboš Pástor and Stambaugh, R. F. (2002), ‘Mutual fund performance and seemingly unrelated assets’, *Journal of Financial Economics* **63**(3), 315–349.

# Appendices

## A Data

### A.1 Variables

**Table A.1:**
**Variable definitions**

This table documents our variables, their definitions, unit of measurement and frequency. We make a distinction between our main variables and other controls.

| Variable | Definition |
|---|---|
| ***Main variables*** | |
| Net flow | Percentage growth in TNA (in USD), net of internal growth (assuming reinvestment of dividends and distributions). See Section 3.2, Equation (6) for details. |
| Excess fee | Total cost of active management, fund fee less comparable index fee. See Section 3.2 for details. |
| Active share | Percentage of a fund's portfolio holdings that differ from its benchmark index holdings. |
| Active fee | Unit cost of active management, excess fee over activity level. See Section 3.2, Equation (8) for details. |
| ***Other controls*** | |
| TER | Yearly expense ratio. Not used directly in any tests. |
| Benchmark-adjusted return (gross alpha) | Difference between the fund gross return and its benchmark return. |
| Fund age | Number of years since the fund's launch date. Used in log form in panel regressions. |
| TNA | Total net assets in millions of U.S. dollars. Used in log form in panel regressions. |
| Family TNA | Family total net assets in millions of U.S. dollars of other equity funds in the same management company excluding the own fund TNA. Used in log form in panel regressions. |
| GDP per capita | Gross domestic product per capita in U.S. dollars in the fund's country (World Development Indicators). Used in log form in panel regressions. |
| Education | Average number of years of education averaged for men and women (World Development Indicators). |
| Fund industry competition | 1 - HHI, where HHI is the sum of squared market shares of fund management companies for mutual funds in the fund's country. |
| Fund industry size | Sum of TNA for all funds within each domicile-year. Used in log form in panel regressions. |
| Market share index | The percentage market share of passive investment vehicles within domicile-years, as measured by percentage of domicile-year TNA. |
| Morningstar rating | A purely quantitative, backward-looking measure of a fund's past performance, measured from one to five stars. Star ratings are calculated at the end of every month. See Morningstar Ratings for more information about the Morningstar ratings. |
| Morningstar style-box | A style based measure based on the tilts toward size (Small, Mid-Cap, or Large) and value tilt (Value, Blend, or Growth) of the fund's actual portfolio holdings. See Morningstar Style-Box for more information about the Morningstar style-box. |

# B Active fee-flow relation in the cross-section

In this appendix, we provide additional analyses on the active fee-flow relation in the cross-section. As laid out in Section 2, our model focuses on the time-varying relationship between net flow and active fee. Furthermore, with a global sample of mutual funds, there is less reason to expect this relationship to be strong in the cross-section than in the time-series due to a large number of potential confounding factors affecting the cross-sectional relationship.

## B.1 Sorting

To understand the active fee-flow relation, we first start by examining the flow-relationship to active share, excess fee, and active fee by single and double sorts. First, we form quintile portfolios based on these variables. Then, we double sort net flow on both excess fee and active share. Then, to get a more granular view than quintile portfolios allow for, we create percentile portfolios and plot active fee, excess fee and active share at time $t-1$ against net flows at time $t$.

The quintiles are constructed within domicile-year pairs based on lagged sorting variables, flows are measured over the ensuing year and are presented in Table B.1. We find that net flows are higher (less negative) in the funds with higher active share (quintile 1 highest and quintile 5 lowest). The same picture emerges for portfolios based on excess fee (quintile 1 has lowest fees). The funds in Q1 had, on average, an outflow of 4.19 %, whereas the most expensive funds (Q5) saw an average outflow of 6.52 %. For active fee, the group with the lowest fee level also saw the least outflows and the most expensive quintile the largest. For all variables (with the exception of Q1 to Q2 for active share), there is a strict negative linear relationship between less favorable variable levels and subsequent flows. Furthermore, the differences between Q1 and Q5 are statistically significant for all. In isolation, both active share and excess fee show the same pattern but combined through active fee the relationship is the strongest with the largest increase in outflows from Q1 to Q5.

**Table B.1:**
**Sorting: Quintiles**

Net flows in quintiles based on active share, excess fee, and active fee. The variables used for sorting is denoted at the top of the table. For active share, Q1 contains the funds with the highest active share, and Q5 the funds with lowest active share. For excess fee and active fee, Q1 contains the funds with lowest fee, and Q5 the funds with highest fee.

| | Net flow | | |
|---|---|---|---|
| Quintile | Active share (H to L) | Excess fee (L to H) | Active fee (L to H) |
| (1) | -4.63 | -4.19 | -3.85 |
| (2) | -4.24 | -5.07 | -4.33 |
| (3) | -5.13 | -5.64 | -5.78 |
| (4) | -6.30 | -6.25 | -6.37 |
| (5) | -7.38 | -6.52 | -7.36 |
| (1) - (5) | 2.75*** | 2.33*** | 3.51*** |

Similarly to Cremers and Curtis (2016), we double sort net flow on active share and excess fee and present these results in Table B.2. This allows us to assess in more detail the relationship between net flows and the intersection of the main components making up our variable of interest. We saw a clear and negative relationship between net flow and quintile portfolios using univariate sorts in Table B.1. As was the case there, here we also witness a more or less negative linear relationship between net flows and decreasing active share within quintiles of excess fee. Keeping the active share variable constant and moving from smallest to largest fee levels, we see for each level of active share that this negative relationship is strictly decreasing. The only exception is when moving from fee category M to L for active share category S. Likewise when keeping the fee categories constant and moving along the active share dimension, much of the same dynamics apply. However, it seems to be for all levels of excess fee a break in the relation when moving from the XL to L category of active share. These results give additional support to the notion that investors appear to take active share into account when selecting funds.

**Table B.2:**
**Double sorting flows**

Net flows in quintiles based on active share and excess fee. Here, we double sort into quintile pairs and calculate average net flows. The value in the top-left cell is the average net flow to the group of funds with the highest active share and the lowest excess management fee.

| | Active share | | | | |
|---|---|---|---|---|---|
| Excess fee | XL | L | M | S | XS |
| XS | -3.31 | -1.51 | -2.43 | -6.45 | -5.64 |
| S | -3.64 | -3.27 | -5.09 | -5.93 | -6.64 |
| M | -4.41 | -3.53 | -5.36 | -6.80 | -8.10 |
| L | -5.49 | -5.69 | -6.12 | -5.59 | -9.02 |
| XL | -5.48 | -6.00 | -6.56 | -6.86 | -9.06 |

To allow for a more detailed picture of the cross-sectional relationship between net flow and active fee, we perform the same exercise but separate funds into percentile buckets and plot active fee at $t-1$ against net flow at $t$. This is shown in Figure B.1, where we also superimpose a linear fit. Figures B.3 and B.2 shows the same, but for the main components that go into our active fee calculations.

It is evident from Figure B.1 that the cross-sectional relationship is negative as well, indicating that active fee is able to explain some of the variations in the cross-sectional flows in the capital in and out of actively managed mutual funds. We infer from this that, on average, investors withdraw capital from actively managed funds that charge a high fee for active management.

**Figure B.1:**
**Active fee-flow relationship**

This plot shows the relationship between average net flow at time $t$ and active fee at time $t-1$. For each domicile-year we rank actively managed funds according to level of active fee (1 = lowest and 100 = highest). Then, we plot the average net flow within each percentile against the average active fee within each percentile. The correlation between $\mathrm{NF}_{i,t-1}$ and $\mathrm{AF}_{t-1}$ is -0.68.

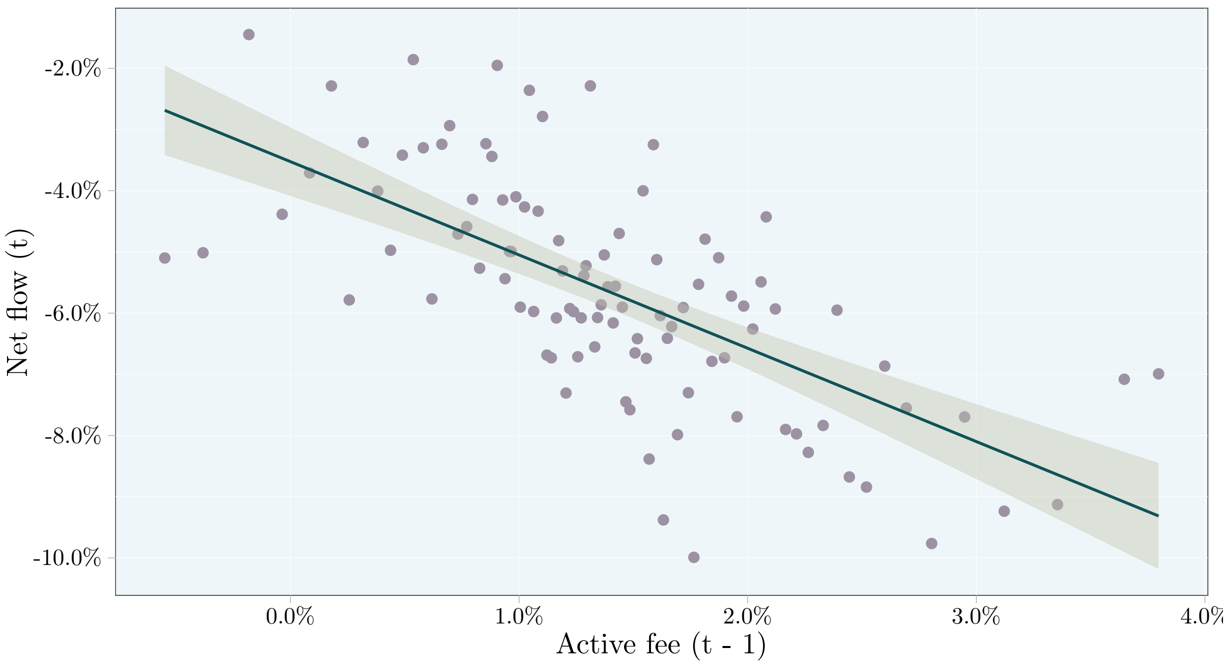


**Figure B.2:**
**Active share-flow relationship**

This plot shows the relationship between average net flow at time $t$ and active share at time $t-1$. For each domicile-year we rank actively managed funds according to level of active share (1 = lowest and 100 = highest). Then, we plot the average net flow against the average active share within each percentile. The correlation between $\text{NF}_{i,t-1}$ and $\text{Active share}_{t-1}$ is 0.56.

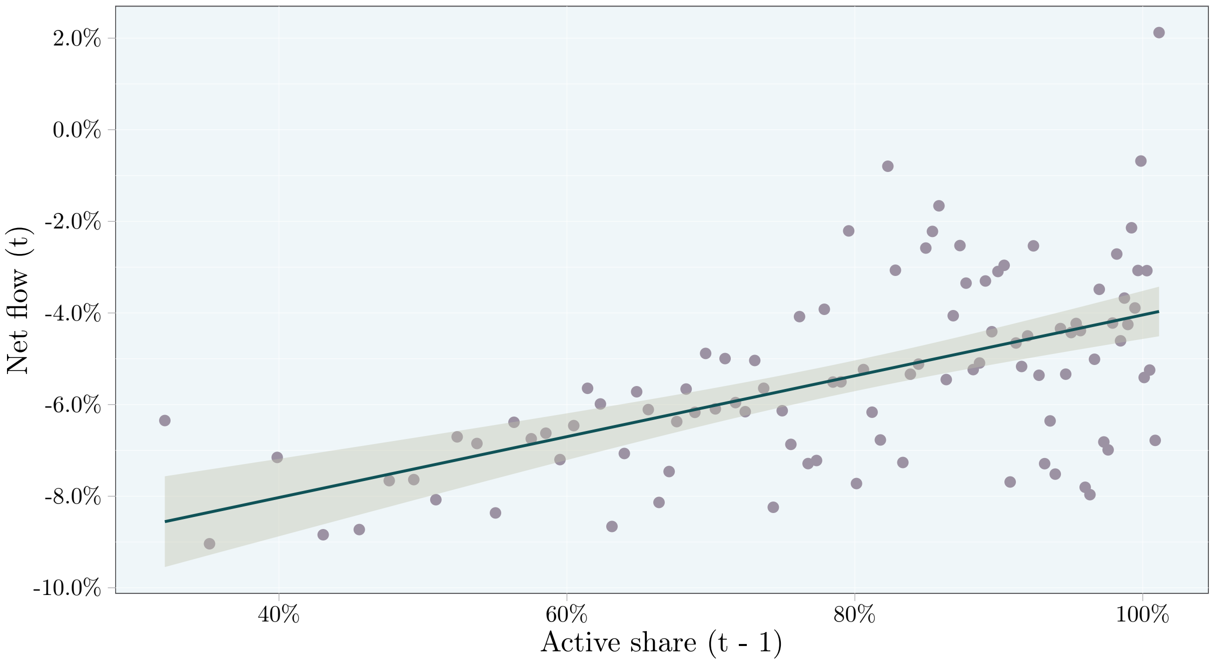


**Figure B.3:**
**Excess fee-flow relationship**

This plot shows the relationship between average net flow at time $t$ and excess fee (EF) at time $t-1$. For each domicile-year we rank actively managed funds according to level of EF (1 = lowest and 100 = highest). Then, we plot the average net flow against the average EF within each percentile. The correlation between $\text{NF}_{i,t-1}$ and $\text{Excess fee}_{t-1}$ is -0.46.