

Forced to be active:
Evidence from a regulation intervention*

Petter Bjerksund Trond Døskeland
André Wattø Sjuve Andreas Ørpetveit

April 29, 2022

Abstract

Closet indexers are low-activity mutual funds that are sold and marketed as active. Their investors therefore only receive part of the service they pay for. Supervisory authorities all over the world are now considering how these funds should be regulated. We examine evidence from interventions carried out by Scandinavian regulators. The impact is identified by comparing scrutinized Scandinavian closet index funds with similar unaffected European funds. Given the choice between reducing fees or increasing activity, the scrutinized funds opt for the latter. Although this results in a more actively managed fund, performance deteriorates. Thus, regulation leads to the worst outcome.

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 this 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 these types of funds are typically looking for an opportunity to outperform the market. To that end, they pay a premium over index funds for the service of dedicated fund managers who try to beat the market.¹ The fund managers' efforts 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 \(2016\)](#)). These funds are labeled "closet indexers". Thus, investors in these funds do not receive the active portfolio management service they pay for.

Financial supervisory authorities around the world have addressed the problem of closet index funds.² The Scandinavian countries (Denmark, Norway, and Sweden) were early to put pressure on potential closet indexers, and carried out extensive investigations in 2014 and 2015.³ 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 exposed to scrutiny with unaffected European closet index funds.

We hypothesize that funds under scrutiny will opt to increase activity and maintain their fee level. Cutting fees is expensive, while increasing activity is cheap by comparison. Rephrased, the funds are "forced" to increase activity. Furthermore, we investigate the consequences of increased activity for investors. One constraint much discussed in the active management literature is diseconomies of scale, i.e., that the quality of marginal investment opportunities declines as active assets under management (AUM) increase. When not subject to an 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 the threshold alpha is zero ([Berk and Green, 2004](#)). However, managers may use different threshold alphas. If they are risk- or effort-averse, they will choose a positive threshold alpha, while if they are overconfident, risk-seeking (e.g., due to convex incentives), or desire to signal skill via activity, they will choose a negative threshold alpha.⁴

Using regulation to induce additional active management may benefit investors if the threshold is positive, but not if it is negative. Our empirical work addresses this by analyzing the effect on investor returns of inducing additional active management. If managers have already exploited their best investment ideas, more activity will lead to lower excess returns. Conversely, a positive threshold

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

²See, for example, [ESMA \(2016\)](#), [New York Office of the Attorney General \(2018\)](#), [Financial Conduct Authority \(2019\)](#), and [ESMA \(2020\)](#).

³See, for example, [Reuters article \(2017\)](#) or [Kjørven \(2019\)](#).

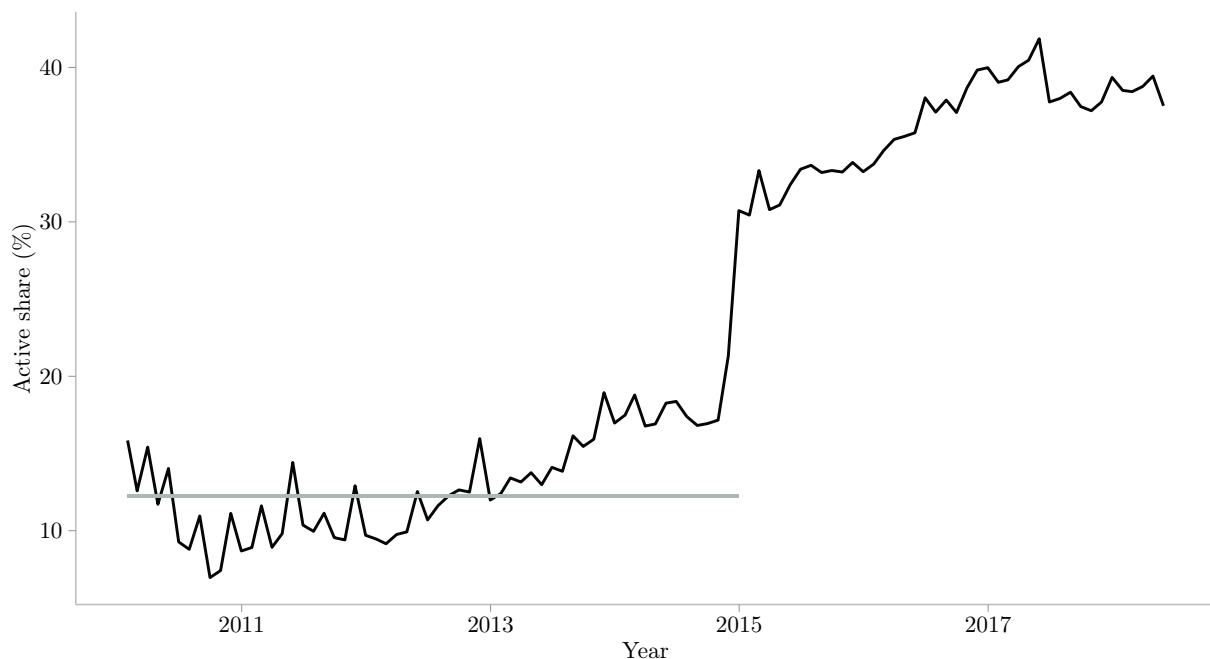
⁴What is optimal for the investor is not necessarily optimal for the fund manager. There are many reasons for conflict of interest between manager and investor. For example, they may have different skills or information sets (see, e.g., [Sirri and Tufano \(1998\)](#) and [Agarwal, Gay and Ling \(2014\)](#)), risk preferences (see, e.g., [Huang, Sialm and Zhang \(2011\)](#)), or incentives (see, e.g., [Sensoy \(2009\)](#) and [Ferreira, Matos and Pires \(2018\)](#)).

alpha implies that investors will benefit from an increased activity level.

Before we report our findings, we set out the history of closet indexing in Scandinavia, focusing in particular on the world's first closet index fund to be ordered by the courts to repay its investors.⁵ Closet indexing has been a concern in all the Scandinavian countries, and the region's different national level Financial Supervisory Authorities (FSAs) have identified questionable practices ([Kjørven, 2019](#)). As an example, 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 DNB Norge's active share from the start of 2010 until May 2018. The Norwegian FSA investigated DNB Norge from 2010 until 2014. A quantitative analysis of the degree of active management identified a tracking error of 1.28% per year and an average active share of 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 development of active share for the mutual fund DNB Norge over the period 2010 to 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 They could bring the fund's management in line with the active management characteristics reflected by the fee and the fund's prospectus or adjust the fee to a level in line with the applied investment strategy. DNB decided to implement a combination of the two. They reduced the fee from 1.8% to 1.4% and

⁵There are some cases where the asset managers have accepted a fine ([Portfolio Adviser article \(2019\)](#)) or where there are ongoing lawsuits ([Financial Times \(Nov. 2021\)](#)).

⁶This corrective order is available with an English translation at [Finanstilsynet](#).

more than doubled the fund's active share. Figure 1 shows how the active share more than doubled by 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 claiming that DNB had misled and charged its investors excessive management fees.⁷ In the first instance, Oslo District Court, DNB won. In the second instance, Borgarting Court of Appeal ruled in favor of the consumers, and ordered DNB to pay back 0.8% of the management fee. DNB then appealed to the Norwegian Supreme Court, whose ruling was delivered on February 27, 2020, upholding the Court of Appeal's judgment. The verdict can be found [here](#). Further details about the lawsuit can be found in Appendix A. The DNB case is unique in that it has brought the closet index problem to court. 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, in that many funds changed behavior to avoid being classified as a closet index fund. The consequences of this new behavior are largely unknown. Closet index funds have become more active, but has this benefited the investors?

At one level, the impact of the intervention is by its very nature interesting. We attribute the changes in fund behavior to the intervention, however we cannot be certain whether the effects can solely be assigned to this or whether they may be due, for instance, to a general change in the conditions for active management. It could be that the timing of the attention from the FSAs coincided with unfavorable market conditions for active management, which means that we cannot definitively conclude what causes a potential low alpha in active management.⁸ These endogeneity concerns must be addressed.

To address these endogeneity concerns and identify the impact of the intervention, we use the exogenous variation created by the policy scrutiny. 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 exposed to scrutiny (treated funds) with European closet index funds not under scrutiny (control funds). In robustness tests, found in Appendix G, truly active Scandinavian funds serve as our control group. Despite the selection problem bias associated with these funds, it is reassuring that the results are largely consistent across control group specification.

We split the impact assessment into two "stages" based on the timing of the event. First, we investigate the impact of the policy scrutiny on active share and fee. In the second step, we examine the consequences for investors' alpha.

Regarding the first stage, we find the unsurprising result that fund managers respond to scrutiny by opting to increase activity over reducing the fee. The diff-in-diff estimate shows an increase in active share of over 9%. Fee estimates are indifferent from zero. For the second stage, we find that the

⁷For details see [Norwegian Consumer Council \(2015\)](#), and [Kjørven \(2019\)](#).

⁸There are different ways of measuring value creation in active management. Most often, we use the term alpha. By that, we mean the difference between the actual portfolio and the benchmark return. This measure is also known as the active or excess return. Alpha can also be used to describe active returns risk-adjusted for factor risk. Finally, alpha can be calculated before or after fees, i.e., gross or net alpha, respectively.

closet index funds that become more active perform worse than a comparable control group of closet index funds. We find that the funds under scrutiny had, on average, a lower annual alpha of 1.3% than comparable funds unaffected by scrutiny.

The empirical analysis shows that scrutinizing closet index funds produces unwillingly active funds, which is the worst alternative. The best solution is not to force funds to become more active but to promote fee reduction. If the regulatory authorities in the DNB case complied with the alternative to update investor information and reduce fees, the investors would have fared better. Assuming the fund had performed as well as the average control fund and had reduced the fee by half of what DNB had to pay back, the annual alpha would have been 1.7% ($= 1.3\% + 0.4\%$) higher.⁹

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](#)). In Section 3: Problem diagnosis, we develop an understanding of the problems associated with closet indexing. Section 4: Intervention design describes the intervention and presents the data with summary statistics. In Section 5: Impact assessment, we examine how fund managers respond to scrutiny and how it affects investors. In Section 6, we go into more detail and examine the robustness of our previous findings. We have tried to keep the number of analyses 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. Although [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 the subject of detailed study and media attention. Many papers have documented that closet indexing exists all over the world ([Cremers et al., 2016](#)).¹⁰ However, there is a difference between identifying the problem and solving it. Market failure often calls for regulation, however it is not always the best solution and may, in fact, make matters worse. There is little literature on the regulation of closet index funds; one exception is [Cremers and Quinn \(2016\)](#), and even less literature examining such regulation. Hopefully, our assessment will 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 to regulators attempting to assist them.

⁹The estimate of loss comes from adding the diff-in-diff estimate for alpha and half the 0.8% DNB had to pay back to its investors.

¹⁰Most of the literature focuses on the connection between active share and performance, see, for example, [Cremers and Petajisto \(2009\)](#), [Cremers, Fulkerson and Riley \(2021\)](#), and [ESMA \(2020\)](#).

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 are unable to manage their financial decisions, they make mistakes that reduce their welfare with broader consequences for the economy. However, regulation can fail and instead exacerbate the problem it set out to solve. As a rule, it is hard to perform a cost-benefit analysis in financial regulation (Sunstein, 2015). Our study provides insight into how the effects of intervention can be assessed.¹¹

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 its raw skill in identifying investment opportunities but also on the various constraints on the fund. A constraint subject to recent discussion is diseconomies of 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 scale effects.

The theoretical model developed by Berk and Green (2004) relies on the key assumption of diseconomies of scale in active management. 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. Table 13 in Adams, Hayunga and Mansi (2021) provides an overview of the literature. 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 relationship between size and performance. However, these findings are challenged by studies that identify an endogeneity issue in the test of the return-to-scale property. Reuter and Zitzewitz (2021) examine the size-performance relationship in a natural experiment setting, applying a regression discontinuity approach. Pástor, Stambaugh and Taylor (2015), Zhu (2018) and Pastor, Stambaugh, Taylor and Zhu (2021) 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) uses 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.

This paper re-examines the size-performance relationship at the fund level with a novel identification strategy addressing the endogeneity issue for domestic European funds. If we accept that asset managers are forced to increase their activity level and have limited opportunities to use in- or outflow to do so, we introduce variation in the quantity of assets under active management. In such a quasi-natural experiment setting, we can identify the size-performance relationship.

Our findings suggest that we have decreasing returns to scale with the mean performance estimates

¹¹Financial Conduct Authority (2016) provides an overview of the British approach to an economic analysis of financial regulation. They outline a methodology for regulatory economic analysis that sets out a three-stage process, including problem diagnosis, intervention design, and impact assessment.

in the treated groups being 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. Our detailed data allow us to learn how managers rearrange their portfolios when pushed to increase their active share. We find that new bets from forced managers underperform new bets from untreated managers.

Finally, a problem associated 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 engage in signal-jamming to appear truly active ([Brown and Davies \(2017\)](#) or [Cremers, Fulkerson and Riley \(2022\)](#)). Signal jamming, i.e., taking random bets, improves closet indexers' chance to pool with genuinely active funds by adding tracking error to their returns, thereby creating a false sense of active management. Moreover, this strategy implies that closet indexing may be more widespread than indicated by traditional measures, such as tracking error and active share.

To distinguish true active management from signal jamming, we examine a portfolio-level concentration measure (used, for example, by [Brands, Brown and Gallagher \(2005\)](#) and [Kacperczyk, Nieuwerburgh and Veldkamp \(2014\)](#)). At the total active portfolio level, we identify more concentrated bets. When dividing the portfolios into sub-portfolios based on bet size, we find that managers take more concentrated bets in new stocks. This finding is in contrast to [Pollet and Wilson \(2008\)](#), who find that funds increase ownership in already owned stocks in response to increased inflows, especially if they operate in relatively illiquid markets. Finally, we find that the increased activity does not come from larger bets in managers' initial "best ideas" ([Pomorski \(2009\)](#) and [Antón, Cohen and Polk \(2021\)](#)).

3 Problem diagnosis

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

3.1 Closet indexing

Closet indexing is problematic because the 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.¹² This information then provides the foundation for the investors' expectation of active management. If this is incorrect, investors will base investment decisions on a false expectation of a more active fund management service than they actually receive

¹²The Key Investor Information Document (KIID) includes the most important investor information. 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.

(Cremers and Quinn, 2016). Closet indexing does not offer the same ex ante risk profile that investors should expect from genuine active management.

Evaluating the services rendered is arduous for mutual fund retail investors because manager effort is only partially observable, and, even if monitoring is possible, interpreting the information may be difficult. Therefore, to mitigate moral hazard, we need suitable measures of effort.¹³ A natural activity measure is the actual outcome of active management, the alpha. 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, alpha is close to zero. Thus, realized returns cannot be used to identify closet index funds. Measures must be based on manager effort.¹⁴

The most used means of identifying closet indexing is active share, which evaluates the degree of active management for funds relative to the benchmark (Cremers and Petajisto, 2009).¹⁵ An alternative is tracking error. We will go on to show that active share and tracking error are highly correlated for European domestic funds. If tracking error is low, active share is low, and vice versa. Therefore, we focus on active share as our activity measure.

To outperform its benchmark, the active portfolio must differ from the benchmark portfolio. As such, activity is a precursor to superior returns. Since closet index funds charge fees akin 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).

3.2 Motives for closet indexing

To crack down on closet indexing, it is important to understand the motives behind it. According to Berk and Green (2004), the quality of marginal investment opportunities declines as active AUM increases. When not subject to an 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 the threshold alpha is zero.¹⁶

However, managers may use different threshold levels. Being risk- or effort-averse leads to a positive threshold alpha. A negative one is chosen if they are overconfident, risk seeking (e.g., due to convex incentives), or have a desire to signal skill via activity. Cremers and Quinn (2016) suggested four

¹³There 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 assume that we can measure the obligation.

¹⁴In the law there is a distinction between the duty to achieve a specific result and the duty of best effort, for more about this, see, for example, UNIDRIOT Principles.

¹⁵Active share is defined as one half of the sum of the absolute value of the difference between portfolio and benchmark weights.

¹⁶With this framework, we relax the assumption that capital is competitively allocated (Berk and Green, 2004). In our setting, there are several reasons why this condition does not hold. Firstly, the event window is short and retail investors are typically slow at changing their positions, i.e., they "suffer" from inertia (see, for example, Biliass, Georgarakos and Haliassos (2010) or Agnew, Balduzzi and Sunden (2003)). Secondly, tax motives do not favor reallocation between funds. Before and during the event, markets went up, meaning that most investors had large and unrealized investment gains. Had they sold, this would have triggered capital gains tax.

alternative motives for closet indexing. We describe two of them, one resulting in a positive threshold alpha and one in a negative one.¹⁷

If managers want to preserve their current asset base and are afraid that underperformance may lead to a large outflow from the fund, they can set a positive threshold alpha. Although a higher active share is optimal for investors, 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 superior returns to grow their asset base, large funds may prefer to preserve their current assets and avoid substantial losses to the benchmark (Sirri and Tufano, 1998). One way to prevent underperformance is to put a larger share of the fund's assets in the benchmark. While this strategy reduces the fund's likelihood of beating the benchmark, average performance may be enough to maintain a large asset pool, and consequently 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 is benched for most of a game. In this case, the investor would be better off if the manager increased the degree of activity.

If managers have run out of new ideas, they may set a threshold equal to or lower than zero. Assuming that managers possess private information about their skills, which are lower than the investors believe, they will run out of good ideas more quickly than set out in the investor information. Size may be a reason why managers run out of ideas. A large fund has a more limited set of investment opportunities consisting of only sufficiently large investments to make a difference. In this case, higher activity does not benefit the investor.

3.3 The rationale for intervening

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, leaving investors worse off than for other investment choices. Thus, there is a potential to increase welfare by accurately regulating these funds.

Cremers and Quinn (2016) suggest two approaches to mitigate the closet indexing problem. Firstly, they suggest a disclosure regime that would incorporate more information, such as active share. Secondly, they examine whether closet indexing is potentially liable under existing laws. In the next section, we show that Scandinavian regulatory authorities used elements of both these approaches during the scrutiny period in our event.

¹⁷We 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 Intervention design

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

4.1 Policy scrutiny in Europe

Although most of the literature focuses on U.S. mutual funds, some papers have investigated European funds (see, for example, [Banegas, Gillen, Timmermann and Wermers \(2013\)](#), [Ferreira et al. \(2013\)](#), [Cremers et al. \(2016\)](#) and [Leippold and Rueegg \(2020\)](#)). Using a difference-in-differences design, we compare Scandinavian funds (treatment group) exposed to policy scrutiny with unaffected funds (control group) in other European countries where we are unable to identify a meaningful level of scrutiny. As well as being the first to have a regulatory focus on closet indexing in Europe (or the world), the Scandinavian countries also have the toughest policy interventions.¹⁸ Below we describe each country's scrutiny in detail.

Denmark

Although we set an exact date in [Table 1](#), these scrutiny processes take a considerable amount of time. 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 limits of 60% for active share and 4% for tracking error, they found that 56 out of 188 equity mutual funds had not practiced the active management strategy they marketed in their prospectuses ([Financial Supervisory Authority 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 was reduced to 22 funds. Based on the report, the FSA contacted the boards of these funds and requested explanations. Apparently, they were satisfied with the answers as no further action was taken. However, the funds were subsequently required to report active share and tracking error.

Norway

During the same period, the Norwegian Financial Supervisory Authority wrote an extensive report on the level of active management for 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 mentioned the case of 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, to 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

¹⁸Two articles from Financial Times: [Financial Times article \(2016a\)](#) or [Financial Times article \(2016b\)](#).

Table 1.
Country selection

This table presents the countries in Europe 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	<i>September 2014</i> Danish FSA publishes a report showing that 56 out of 188 funds studied were potential closet indexers. More details in text.
Norway	<i>March 2015</i> Norwegian FSA issues a corrective order to DNB asset management regarding active management. More details in text.
Sweden	<i>October 2015</i> Finansinspektionen (FI) publishes a report debating stricter rules on consumer protection in financial markets. 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	<i>March 2016</i> The Italian regulator takes action against some of the largest investment companies in its domestic market for mis-selling actively managed funds that closely hugged an index.
Netherlands	<i>May 2016</i> AFM publishes a report on index hugging identifying 7 out of 85 funds investigated as closet trackers.
Germany	<i>September 2016</i> BaFin completes its investigation into closet indexing, identifying deficiencies in transparency.
France	<i>March 2017</i> AMF reminds asset management firms of the importance of clarity in the investment objective.
United Kingdom	<i>June 2017</i> FCA publishes their final report on the Asset Management Market study finding £109 bn invested in closet funds.
Luxembourg	<i>August 2017</i> CSSF issues a reminder on improving clarity in the “objectives and investment policy” section of the KIID.
Spain	<i>October 2018</i> CNMV analyzes the existence of these products without reaching any conclusion on how to act. More details here.
Ireland	<i>July 2019</i> Central Bank of Ireland publishes largest data driven study of industry about closet indexing to date.

fee to a level in line with the strategy applied. The second fund that the FSA criticized was Nordea Avkastning. They were given the same options, either change the level of activity or the fee, although the activity level was higher than for DNB Norge.¹⁹

Sweden

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 provided accurate and clear investment objective and policy information. The investigation is presented in the Swedish FSA's annual report on Consumer Protection 2015 ([Financial Supervisory Authority of Sweden, 2015](#)). The intervention started a debate on the legal issues related to closet indexing. In 2014, the Swedish Shareholders' Association (Sveriges Aktiesparares Riksförbund) initiated a class action against two mutual funds from one of the largest Swedish banks, deciding, however, not to go through with the lawsuit in July 2015.²⁰ For the Scandinavian countries, [Kjørven \(2019\)](#) analyzes how the European legal framework has been applied and discusses the need for legal measures to ensure that investors get what they pay for, as protection against closet indexing.

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 prevalence of closet indexers. Their findings suggest that between 5% and 15% of the sample funds were potential closet indexers. As we see in Table 1, several other countries performed their own investigations. To our knowledge, none of these have resulted in any legal claims against funds.

4.2 Event design

To conduct an event study, we need to follow the mutual funds before, during, and after the intervention. 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 December 2015, 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 subsequent media interest, we assume that the funds at risk of being labeled a closet index fund were rethinking their active management strategy during the two-year window from January 2014 to December 2015.

¹⁹The Financial Supervisory Authority of Norway, available in Norwegian at [Finanstilsynet](#).

²⁰For details about the funds, see case number 2014-11304 ([Swedish Ministry of Finance, 2016](#)).

To identify a control group as free as possible from scrutiny, we have sorted all the countries in Europe, except Scandinavia, into two groups, those with and without identified policy scrutiny. In Panel II: Control in Table 1, we list the countries that constitute our control group, and in Panel III: Omitted, we list the countries that have scrutinized the fund industry. For the latter group, we also add links with further details.

We have also performed a robustness test where we replace our primary control group with truly active funds from Scandinavia. Due to a selection problem, i.e., comparing "sick" closet index funds with "healthy" truly active funds, we do not use this as our primary control group. However, we find almost identical results, indicating robust results. For details, see Appendix G.

4.3 Data and summary statistics

This section describes the data, how we constructed 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 span from January 2010 through December 2019. Our focus is on domestic long-only equity funds.²¹ The fund data include monthly assets under management and gross and net returns.²² The fee is the price of active management and is calculated using the difference between gross and net returns. To calculate active share, we use monthly holdings for each fund.²³ Details on the sample selection and raw data are presented in Appendix B.1.

A benchmark is defined for each fund to calculate active share and performance. In general, categorization is complicated for active funds, however the error is minimized by using domestic funds. These funds give us the most homogeneous group of funds both across and within countries. 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 may not accurately reflect their actual investment style. Details regarding benchmark data are provided in Appendix B.2.

To avoid survivorship bias, we also include funds that died during the sample period. However, we exclude funds with less than 12 months of data in either the pre- or post-event period to draw meaningful inference. After these exclusions, we are left with a sample consisting of 353 funds, from which we define closet index funds. The full sample is presented in Table B.2 in Appendix B.3.

²¹We define domestic funds as funds with an investment area equal to the home country of the investment company and a domestic primary prospectus benchmark.

²²Returns are in the local currency, while assets under management are in USD to establish a common currency for comparison across countries.

²³We have used Morningstar data for funds missing data in Lipper.

To evaluate factor returns, we collect size and style portfolio returns from MSCI. All variables, divided into outcome and controls, can be found in Table B.3 in Appendix B.4. Continuous variables are winsorized at the 1st and 99th percentiles to mitigate the potential impact of outliers.

Defining closet indexing

There are several ways to define a closet index fund, the two most common of which are active share and tracking error. In Appendix B.5, we show that the factor structure is weak for domestic funds, which implies a high correlation between active share and tracking error. Thus, we choose to focus on active share in our analyses. To define a fund as a closet index fund, we set two cutoff points, 40% and 50%, with funds being classified at the start of the event window. We hypothesize that scrutiny exposure increases with lower activity levels. Therefore, when comparing the two samples, the estimates are largest for the lowest cutoff point.

4.4 Summary statistics

In Table 2, we report summary statistics for the treated sample (Scandinavian funds) and the control sample (European, non-Scandinavian funds). Active share and other fund characteristics are measured before the event window. 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 accurately than the loadings on risk factors. However, in robustness tests, factor-adjusted alphas are also analyzed. The variables are explained in Table B.3 in Appendix B.4.

Table 2.
Summary statistics: closet index funds

This table presents summary statistics for the sample of closet index funds in Scandinavia (Norway, Sweden, and Denmark) and the rest of Europe. Values are means over a two-year window before the event start. Gross alpha, expense ratio, and net alpha are annualized. Displayed values for competition are country-means for the Scandinavian countries and the rest of Europe, respectively (see Table B.3 for details). Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. Inference on differences between treatment and control funds is based on Newey and West (1987) standard errors. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Active share $\leq 40\%$			Active share $\leq 50\%$		
	Treated	Control	Difference	Treated	Control	Difference
Number of funds	46	33		75	47	
Active share (%)	32.4	32.2	0.2	38.0	36.5	1.5
Gross alpha (%)	1.04	0.59	0.46	1.24	1.03	0.20
Expense ratio (%)	1.14	1.18	-0.04	1.25	1.19	0.06
Net alpha (%)	-0.13	-0.60	0.47	-0.05	-0.18	0.12
AUM (million USD)	474	221	253**	377	225	152
Fund age (years)	16.2	16.0	0.2	15.0	15.4	-0.4
Competition	0.84	0.81	0.03	0.84	0.81	0.03

In columns 1 to 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 32%. Regarding alpha, we find similar results as other studies in 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. Control variables and fixed effects are used to correct for these differences.

In columns 4 to 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 for the pool of funds is around 37%. Again, the treated and control sample are similar along all dimensions, now also for size. It is worth noting that compared to a typical U.S. fund, the European funds are small in terms of their management teams and organization scope but large relative to their investment space.

5 Impact assessment

We study the effect of the interventions in two "stages". In the first stage, we examine the impact of scrutiny on the level of activity and fees. We have already hypothesized that the funds will increase activity over reducing fees. For the effect on alpha in the second stage, we explore the consequences of the potential new behavior. Depending on the motives for closet indexing, described in Section 3.2, we can expect either a positive or a negative impact on alpha.

Our goal in describing the event is twofold. Firstly, we want to document what really happened to the outcome variables, i.e., active share, fee, and realized alpha. Secondly, we want to understand whether differences in realized outcomes result from noise or can be interpreted as "significant" differences. Therefore, we start by describing the realized effects and then "add on" additional statistical tests.

5.1 First stage: Impact on active share

In this section, we empirically test what happens to active share for funds under scrutiny. Figure 2 illustrates the development of the realized mean active share for treated and control funds over time. We clearly find that funds exposed to scrutiny increase active share more than unaffected funds. We also see that the two groups not only display parallel trends before the event but also the same level of active share.²⁴

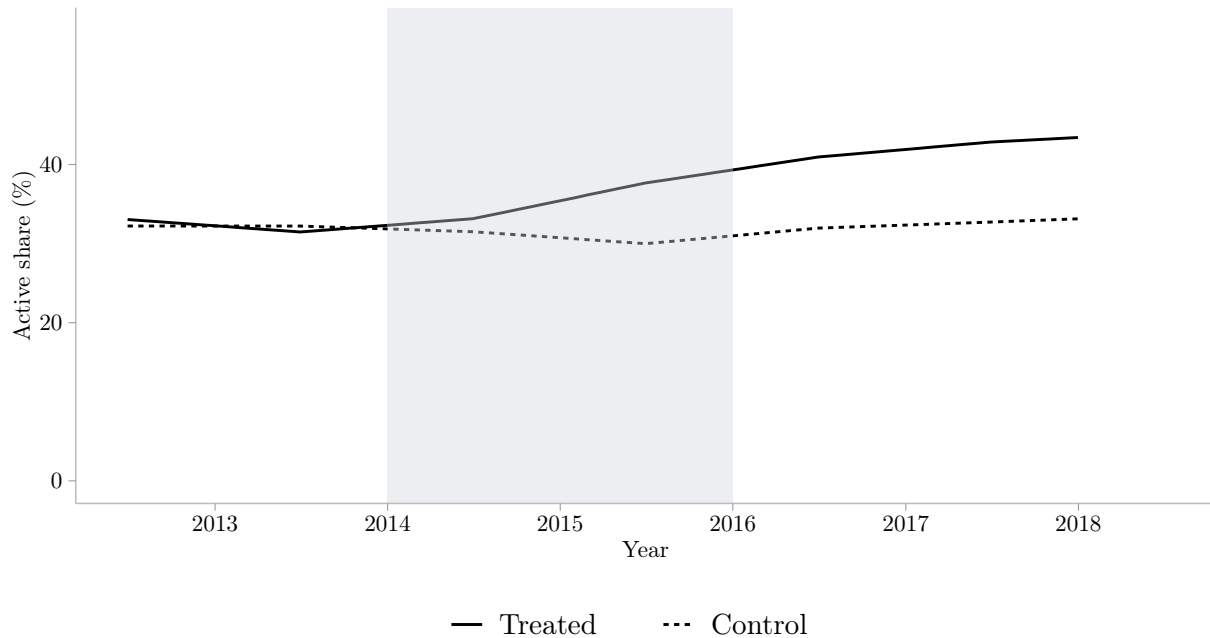
In Table 3, we report the realized differences for both the sample with active share below 40% and 50%. Results are reported using a two-year pre-and post-event window. For the below 40% sample, the difference between treated and control is 9.4% after the event. As shown in Figure 2, the funds' activity is similar before the event. Thus, the diff-in-diff estimate of 9.2% is roughly the same as the post-event estimate. The increase is almost one third ($9.24\%/32.4\%$) relative to the pre-event level. For

²⁴ A 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 development in active share for closet index funds¹ from 2012 through 2017. The time series are the annual averages of monthly cross-sectional, group-wise means, and the shaded area highlights the event window.

¹ Active share $\leq 40\%$ in 2013.



the sample of funds with active share below 50%, the post-event difference is 7.3%, and the diff-in-diff estimate is 5.7%. Thus, we find that scrutiny leads to a higher activity level, with the effect being larger for the most intensely scrutinized funds.

Table 3.
Effect of intervention on active share

This table reports the effects of policy interventions on **active share** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. Estimation is based on two-year pre- and post-event averages. The difference column presents post-event differences between treated and control funds. The difference-in-differences follows a classic setup, see [Angrist and Pischke \(2008\)](#) for details. [Newey and West \(1987\)](#) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Active share	
	Difference	Difference-in-differences
Active share $\leq 40\%$	9.41*** (2.50)	9.24*** (2.29)
Active share $\leq 50\%$	7.28*** (2.27)	5.70*** (2.05)

Regarding statistical inference, we also calculate the p-value for the different estimates in Table 3. In this first statistical analysis, we use the average difference between a treated and non-treated fund for the two years before and 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). We find that all estimates have a p-value below 1%.

By splitting the sample into different groups based on fund characteristics, we examine whether the impact comes from certain funds. We divide the sample into a high or low category based on whether the fund is above or below the median value. The "high" treated funds are then tested against the "high" control funds, and similarly for "low" funds. The estimates from the triple diff-in-diff regressions are shown in Table C.2 in Appendix C. Overall, we find that positive active share estimates are present for all fund categories: size, age, fee level, and past performance. However, the age characteristic is interesting; for both the 40% and 50% sample, the old funds respond significantly more than the young funds. This result may indicate that the older closet index funds are more concerned with scrutiny.

We also formally test the effect of scrutiny using fixed effects (fund and time) panel regressions to minimize endogeneity concerns, among other things. Fixed effects soak up any variation in active share due to cross-sectional differences in fund characteristics, meaning that identification comes from variation within a fund over time, and not from the cross-section of funds. These estimates are presented in Table C.3 in Appendix C. The coefficient sizes and significance levels are similar to the previous specifications, i.e., the estimates are large and significant.²⁵

5.2 First stage: Impact on fee

In this section, we empirically test the effect of scrutiny on fees. We have hypothesized that fund managers opt not to adjust fees. Again, we present two types of results, one documenting what happened and one where we statistically test for effects.

Figure 3 illustrates the development of the mean fee for treated and non-treated funds. 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 minor and stable, meaning that funds under scrutiny did not reduce their fees more than unaffected funds. In Table 4, we report realized differences, given as annualized figures. The estimates show that scrutiny did not lead to a relatively lower fee level. We also check whether the p-value is below the regular level for statistical significance and find that none of the estimates are significantly different from zero.²⁶

Table D.2 in Appendix D shows the sample split into different subsamples based on fund characteristics. We find no meaningful group-wise differences. If anything, scrutinized closet index funds with an active share below 50% and relatively low fees or age increase fees more than the control group. Table D.3 in Appendix D tests the effect of scrutiny using fixed effects (fund and time) panel regressions. Not surprisingly, the estimates are low and insignificantly different from zero.

To sum up for fees, every estimate is small and insignificant, thus confirming our hypothesis that managers under scrutiny choose to increase activity over reducing fees. This result supports the

²⁵The control variables are described in Appendix B.4.

²⁶For a pre-trend analysis, see Appendix D.1.

Figure 3.
Development of fees

This figure presents the development in fund fees for closet index funds¹ from 2012 through 2017. The time series are the annual averages of monthly cross-sectional, group-wise means, and the shaded area highlights the event window.

¹ Active share $\leq 40\%$ in 2013.

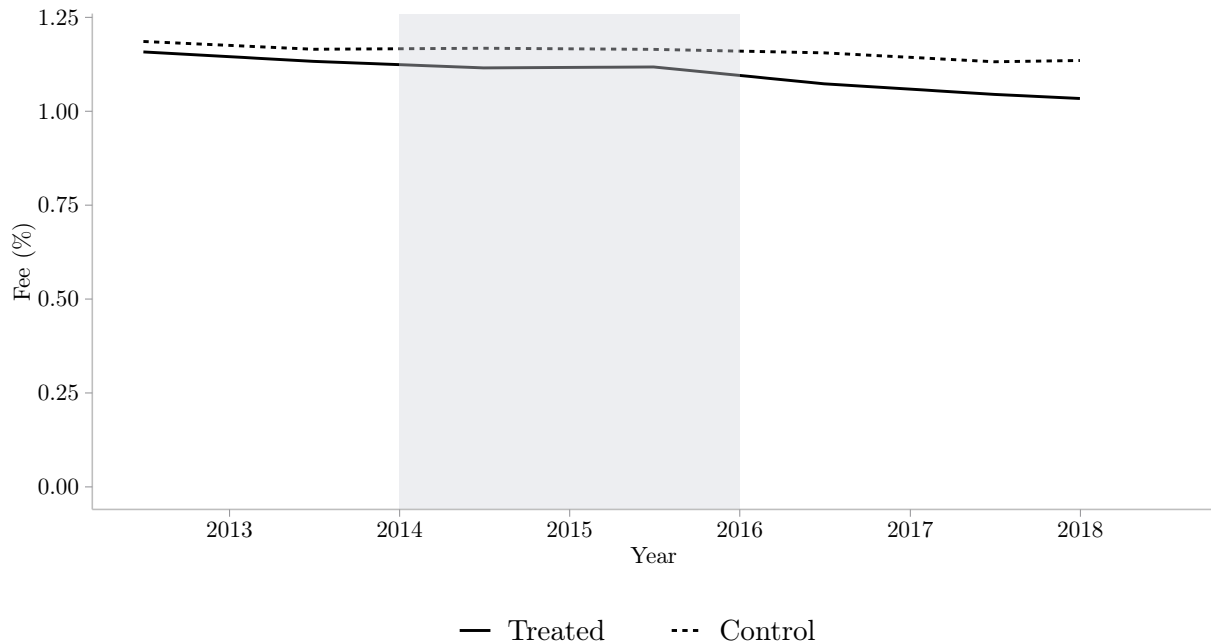


Table 4.
Effect of intervention on fees

This table reports the effects of policy interventions on **fees** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. Estimation is based on two-year pre- and post-event averages. The difference column presents post-event differences between treated and control funds. The difference-in-differences follows a classic setup, see [Angrist and Pischke \(2008\)](#) for details. Reported coefficients are annualized from fund-month level observations. [Newey and West \(1987\)](#) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Fee	
	Difference	Difference-in-differences
Active share $\leq 40\%$	-0.10 (0.12)	-0.07 (0.05)
Active share $\leq 50\%$	0.06 (0.12)	0.00 (0.04)

notion that the managers were "forced" to increase activity. The following section investigates the consequences of increased activity on performance.

5.3 Second stage: Impact on alpha

In this section, we study the total impact of policy scrutiny on fund performance. As mentioned above, arguments exist for both a positive and negative impact on performance. Our dependent variable is

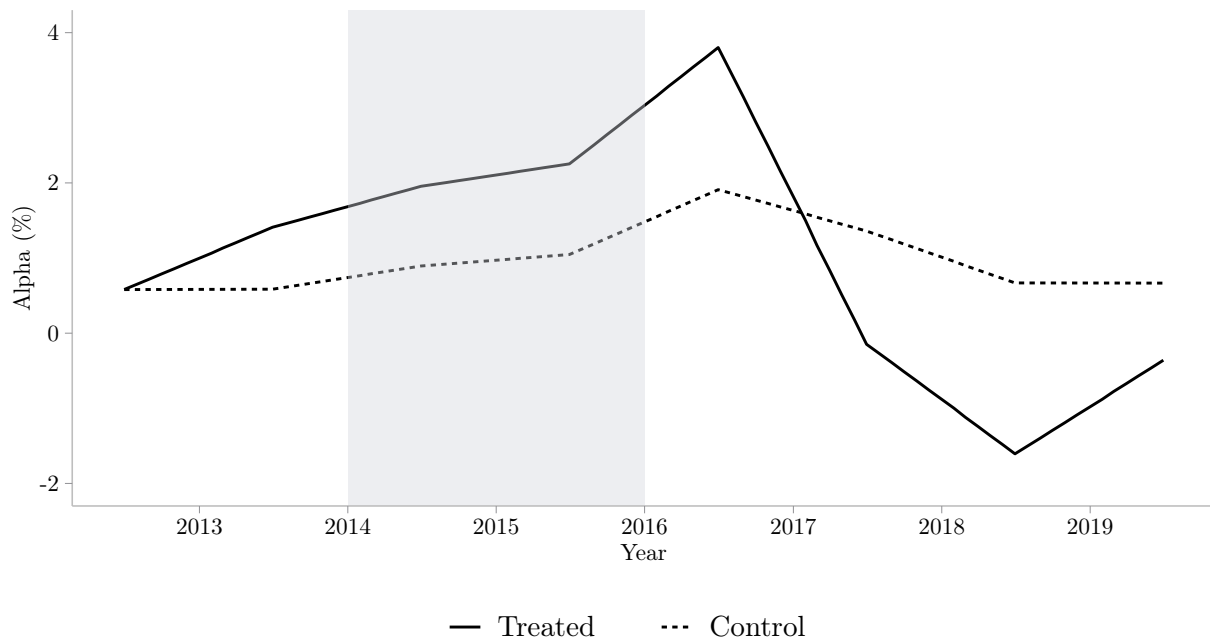
alpha, and, as demonstrated above, the changes in active share have already taken place going into the post-event period. To evaluate the impact, we use a four-year post-event period. Again, we first calculate the actual effect, i.e., how much return the group of scrutinized funds has generated relative to the group of unaffected funds, before formally testing whether this impact is statistically significantly different from zero.

Figure 4 illustrates the development of realized mean gross alpha for treated and non-treated closet index funds. In the pre-event period, treated funds appear to perform somewhat better than non-treated funds. However, this trend is insignificant, as reported in Table E.1 in Appendix E. After the intervention, we see a different pattern; treated funds underperform the control funds.

Figure 4.
Development of alpha

This figure presents the development in fund alpha for closet index funds¹ from 2012 through 2019. The time series are the annual averages of monthly cross-sectional, group-wise means, and the shaded area highlights the event window.

¹ Active share $\leq 40\%$ in 2013.



In Table 5, we report the realized differences. The annual difference between funds below 40% active share under scrutiny and not under scrutiny is 0.83% after the event. Since treated funds performed better before the event, the diff-in-diff estimate is 1.29%. 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 the intervention would have lost 1.29% annually. Given the lower increase in active share for the 50% cutoff, we expect the results to be less negative. Surprisingly, we find similar point estimates for both samples.

Regarding statistical inference, there are many issues related to performance evaluation and there is no academic consensus on a particular method of testing performance (see, for example, Elton and

Gruber (2020) for a recent review of the relevant literature). Therefore, we emphasize the importance of using alternative approaches.

Table 5.
Effect of intervention on alpha

This table reports the effects of policy interventions on **alpha** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. Estimation is based on two-year pre- and four-year post-event averages. The difference column presents post-event differences between treated and control funds. The difference-in-differences follows a classic setup, see Angrist and Pischke (2008) for details. Reported coefficients are annualized from fund-month level observations. Newey and West (1987) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Alpha	
	Difference	Difference-in-differences
Active share $\leq 40\%$	-0.83** (0.34)	-1.29** (0.58)
Active share $\leq 50\%$	-1.10*** (0.34)	-1.30*** (0.50)

Firstly, we calculate the p-value for the different estimates in Table 5, and find that they all have a p-value below 5%.

Secondly, we perform a placebo test. One potential concern is that the estimated impact of scrutiny is either 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 by means of a placebo test where we randomly assign funds to treatment and control groups, maintaining the same ratio of treated to non-treated funds as in the original sample (see Table 2).

Using these randomly assigned groups, we estimate the diff-in-diff model presented in Table 5. We repeat this exercise for 1,000 estimations and report the results in a histogram in Figure E.1 in Appendix E. We find a significantly negative effect (5% confidence level) for only 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 coefficients are equal to or smaller than the coefficient estimated using the original sample, represented by dashed vertical lines in Figure E.1. These results reassure us that our tests capture the treatment effect of regulatory scrutiny on fund alpha and not some random effect or omitted variable.

Thirdly, we split the sample into different groups based on fund characteristics. We examine whether the impact stems from certain types of funds. The triple diff-in-diff regression estimates are shown in Table E.2 in Appendix E. Overall, estimates are negative for all fund categories: size, age, fee level, and past performance. However, they do not all have a p-value below 5%. As we split the

sample however, the number of observations decreases and the variance of the estimator increases, making it harder to achieve statistical significance.

Fourthly, in Table 6 we perform fixed effects (fund and time) 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 related to time. We estimate the difference in alpha between funds exposed and not exposed to scrutiny. In the first two columns, we do not add controls, while in 3 and 4, we control for fund age, size, and fees. We find that the estimates are about the same as using standard diff-in-diff regressions.

Table 6.
Alpha-scrutiny relationship

This table reports estimated slope coefficients ($\text{Post} \times \text{Scrutiny}$) from panel regressions testing the effects of policy interventions on **alpha** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All regressions include fund and month fixed effects, and columns 3 and 4 additionally control for fund age, size, and fees. Reported coefficients are annualized from fund-month level observations. Standard errors (in parentheses) clustered by fund are reported in columns 1 and 3 and independently by fund and month in columns 2 and 4. There are 5,492 and 8,410 fund-month observations in the 40% and 50% panels, respectively. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Post \times Scrutiny			
	(1)	(2)	(3)	(4)
Active share $\leq 40\%$	-1.13** (0.49)	-1.13 (1.16)	-0.93* (0.51)	-0.93 (1.23)
Active share $\leq 50\%$	-1.12** (0.43)	-1.12 (1.31)	-1.07** (0.45)	-1.07 (1.36)
Controls			\times	\times
Fund cluster	\times	\times	\times	\times
Month cluster		\times		\times
Adj. R ² [40% 50%]	[0.10 0.13]		[0.10 0.13]	

Regarding p-values, there are differences between the specifications in Table 6. In all specifications, standard errors are clustered along the fund dimension, however in columns 2 and 4 we independently cluster along the time dimension as well. This two-way clustering sharply increases the estimated standard errors. This is probably due to an intra-time dependence structure in mutual fund returns that is present even after time fixed effects are removed.²⁷

As a final robustness test, we use two alternative measures of value creation, i.e., net alpha and factor-adjusted alpha. So far, we have assumed that the effect on net and gross alpha is the same. This assumption is based on documented small fee effects. Columns 1 and 2 in Table E.3 in Appendix

²⁷ 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. For example, see the recent well-published papers Pástor et al. (2015), Pástor and Vorsatz (2020), and Cremers et al. (2021).

E confirm our assumption that the results also hold when net alpha is the dependent variable.

Limiting our scope to comparing European domestic funds means we do not expect any funds to have a particular factor style, since the high correlation between active share and tracking error indicates a weak factor structure. Funds with a factor style often have a broader investment universe than a single European country. Consequently, our focus has been on adjusting for risk prescribed by the benchmark. However, as a robustness test, we adjust alpha for domestic factor risk and perform the same regression analyses as in the other tables. Columns 3 and 4 in Table E.3 show that the estimates have the same sign but slightly lower magnitude after adjusting for factor exposure. Thus, the negative alphas are caused by a combination of poor stock picking and factor exposure.

To sum up, our analyses show that scrutinized closet index funds underperform non-scrutinized closet index funds. These findings relate to our suggested motives for closet indexing in that they support the notion that the manager runs out of ideas for new successful bets. Thus, regulation then forces managers to take bets in a sub-optimal manner, leading them to destroy investor value. Taking this into consideration, the best alternative for investors, conditional on supervisory intervention, is reduced fees.

6 Further analyses

Thus far, we have shown that scrutinized closet index funds increased activity and achieved lower alphas than non-scrutinized funds. In Section 6.1, we decompose the total active portfolios into sub-portfolios. Furthermore, we examine the source of increased active share and how this impacts performance. We argue that this detailed analysis can also be viewed as a test of diseconomies of scale in active management. Finally, in Section 6.2, we introduce an additional activity measure to determine whether the change in active share is a result of true active management or just signal jamming.

6.1 Decomposing the active portfolio

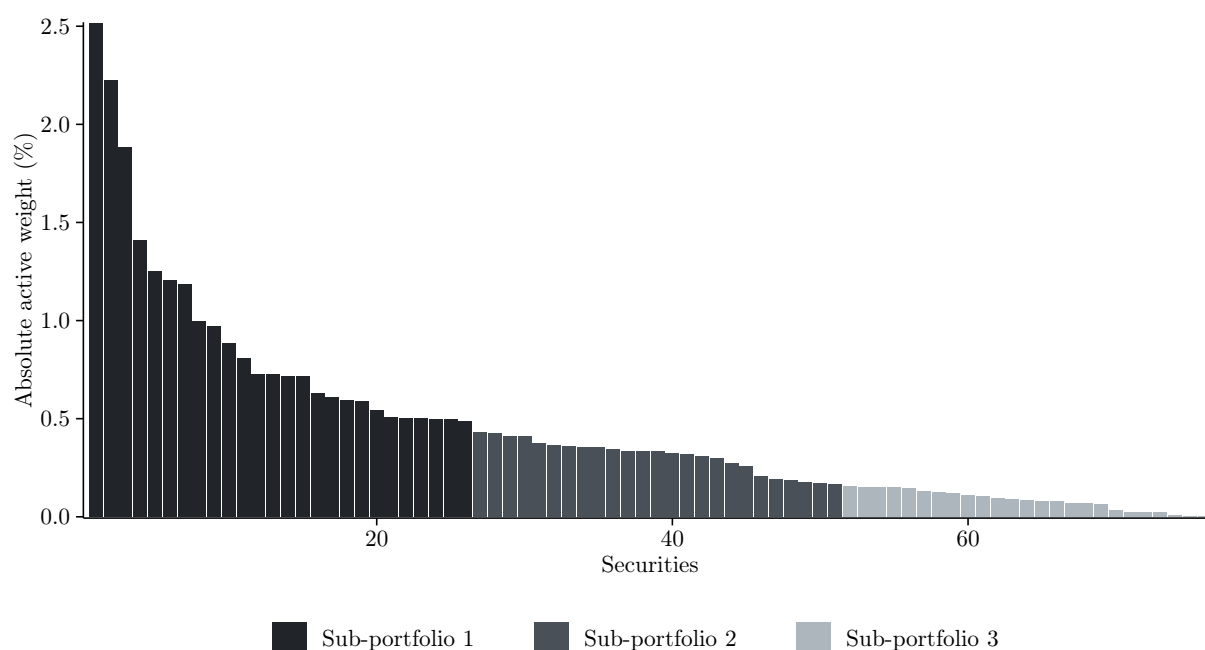
The nature of our data allows us to split the total active portfolio into subgroups, which is done based on active position size. This gives us three sub-portfolios per fund containing the largest, middle, and smallest positions based on absolute active weights. Inspired by Pomorski (2009) and Antón et al. (2021), we label the portfolio containing the largest bets the "best ideas" portfolio. To illustrate this categorization, Figure 5 shows how we allocate all the bets to a corresponding sub-portfolio based on all active weights for DNB Norge at the end of December 2013 (end of pre-event period).

We also identify a sub-portfolio including all the new bets obtained during the policy scrutiny period, which we label "new bets". These bets can come from existing stocks without an initial bet or from newly listed stocks.

In Table 7, we present our findings for the four sub-portfolios. Regarding changes in active share,

Figure 5.
Active weights for DNB Norge in December 2013

This figure presents 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.



we find that the difference between treated and control is positive for both the new and the smallest stock portfolio.²⁸ This applies to both the group of funds with an active share below 40% and 50%.²⁹

Furthermore, we find that new bets taken by scrutinized funds underperform those taken by non scrutinized funds. This is contingent on the main source of scrutinized funds' portfolio underperformance set out in Section 5.3 stemming from these new bets.

In total, relative to active managers not scrutinized, treated funds generate roughly the same alpha on their best ideas and bets that they are already familiar with but lose out on new stocks. We can only speculate, but one potential explanation is that closet index managers are uncomfortable with being forced to increase activity. This result underpins a lack of new quality bets being the motive for closet indexing.

Diseconomies of scale in active management?

One constraint discussed in the active management literature is diseconomies of scale. If scale impacts performance, skill and scale interact. A more skilled large fund can in such instance underperform a

²⁸Note that the sum of the characteristics we measure for the sub-portfolios is not identical to the total portfolio. This is because our definition of the sub-portfolios is limited to capturing the changes in stocks listed in December 2013 and does not capture all newly listed companies after the event window.

²⁹All the funds follow the UCITS directives. The best known restriction is the so-called "5/10/40 rule". In summary, it sets out that a maximum of 10% of a UCITS fund's net assets may be invested in stocks from a single issuer and that investments of more than 5% with a single issuer may not make up more than 40% of the whole portfolio. One could argue due to this restriction; treated funds cannot increase their already largest bets. However, this is not likely since the restriction should also be binding for control funds.

Table 7.
Effect of intervention on sub-portfolios

This table reports the effects of policy interventions on sub-portfolio **active share** and **alpha** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All results are from a difference-in-differences model equivalent to the one presented in column 2 in Tables 3, 4, and 5. For active share, estimates are based on two-year pre- and post-event averages, while alpha estimates are based on two-year pre-event and four-year post-event averages. Reported coefficients for alpha are annualized from fund-month level observations. Newey and West (1987) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Active share				Alpha			
	New stocks	SP 1	SP 2	SP 3	New stocks	SP 1	SP 2	SP 3
Active share $\leq 40\%$	3.28*** (0.44)	0.60 (1.35)	0.30 (0.48)	0.72** (0.35)	-0.44*** (0.13)	0.17 (0.34)	-0.18 (0.16)	0.16 (0.12)
Active share $\leq 50\%$	3.10*** (0.40)	-0.70 (1.05)	-0.08 (0.43)	0.83*** (0.30)	-0.55** (0.23)	0.25 (0.38)	-0.34** (0.15)	-0.12 (0.33)

less skilled small fund. Consequently, we must understand scale effects to learn about skill.

Under certain assumptions, the intervention we examine enables us to test diseconomies of scale. If we assume that asset managers are forced to increase their level of activity and cannot use in- or outflows to change the degree of active management, we get a quasi-natural experiment that allows us to identify the scale-performance relationship. Given the diseconomies of scale predictions, described in Section 2, and the findings so far, we have indications of this type of dynamics.

6.2 Signal jamming?

As described in Section 2, one potential concern is that the change in activity does not come from real bets but from managers adding noise to their returns through signal jamming (Doshi, Elkamhi and Simutin (2015) and Brown and Davies (2017)). A problem with the activity measures used so far is that they cannot detect whether funds engage in signal jamming to appear truly active. Therefore, we introduce a measure of stock-level portfolio concentration. This measure is used at the stock-level by Brands et al. (2005) and Kacperczyk et al. (2014), and at the industry-level by Kacperczyk, Sialm and Zheng (2005). The portfolio concentration measure is calculated by summarizing the squared active weights, meaning that the measure places more weight on larger active positions.³⁰

By comparing the development in portfolio concentration for treated and 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 portfolio concentration. Conversely, if more concentrated bets are taken, we expect the diff-in-diff estimate to be positive.

We use the same methodology to describe the development of the active portfolio concentration, as for testing the change in active share. Regression results are presented in Table 8. If we assume the

³⁰Portfolio concentration is defined as $\sum_{i=1}^N (w_i^a)^2$, i.e., the Herfindahl-Hirschman index of the portfolio's active weights.

active portfolios for treatment and control funds are equal before the event, we can see from column 1 that the treated post-event portfolios consist of more concentrated active bets. The difference is only statistically significant for the sample with a cutoff at 40%. The diff-in-diff analysis yields the same results with smaller estimates. Similar to [Kacperczyk et al. \(2005\)](#), it is harder to use portfolio concentration to interpret economic significance than when using active share. In Appendix F, we perform fixed effect regressions and obtain the same results.

Table 8.
Effect of intervention on portfolio concentration

This table reports the effects of policy interventions on **portfolio concentration** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. Estimation is based on two-year pre- and post-event averages. The difference column presents post-event differences between treated and control funds. The difference-in-differences follows a classic setup, see [Angrist and Pischke \(2008\)](#) for details. [Newey and West \(1987\)](#) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Portfolio concentration	
	Difference	Difference-in-differences
Active share $\leq 40\%$	1.16*** (0.26)	0.62*** (0.22)
Active share $\leq 50\%$	0.46 (0.51)	0.33 (0.26)

In Table 9, we perform similar analyses to investigate different parts of the active portfolios. We find that treated managers take more concentrated bets in stocks not previously part of the active portfolio for the cutoff at 40% and 50%. Thus, we find that scrutiny makes funds take more concentrated bets. A significant part of these bets is concentrated in the "new bets" part of the active portfolio.

Table 9.
Effect of intervention on sub-portfolio concentration

This table reports the effects of policy interventions on **sub-portfolio concentration** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All results are from a difference-in-differences model equivalent to the one presented in column 2 in Tables 3, 4, and 5. Estimation is based on two-year pre- and post-event averages. [Newey and West \(1987\)](#) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Portfolio concentration			
	New stocks	SP 1	SP 2	SP 3
Active share $\leq 40\%$	0.10*** (0.03)	-0.04 (0.26)	0.07** (0.03)	0.04* (0.02)
Active share $\leq 50\%$	0.10*** (0.02)	-0.41* (0.21)	0.04 (0.03)	0.04* (0.02)

7 Policy implication and conclusion

The current trend is that people are increasingly becoming their own money managers. The landscape in which they make decisions is becoming more and more complicated, and many often lack the information and knowledge they need 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, financial authorities need to align the interests of financial intermediates and investors.

In this paper, we have attempted to identify what can be learnt from the interventions in the Scandinavian countries. We highlight the following four results. Firstly, a potential closet index fund opts to increase activity and not reduce fees when placed under suspicion by the supervisory authorities. Secondly, when closet index funds are forced to become more active, they perform worse than unaffected comparable funds. Thirdly, the funds under scrutiny take positions in new stocks when they increase activity. Finally, the funds under scrutiny lose out on these new marginal bets relative to unaffected funds, which lends support to the hypothesis on diseconomies of scale in active management.

Regulators face two principal alternatives. They can either force funds to increase the active management level or reduce fees. So far, little research has investigated the effect of this intervention. We show that when given the choice, the funds increase their level of activity. We propose two different rationales for this behavior; managers are either skilled and in possession of new ideas but are afraid of losing assets under management or they lack ideas and want to harvest as much revenue as possible. Activity-increasing managers' failure to create value for their investors supports the narrative of revenue harvesting managers without a plethora of new, good ideas.

What can regulators thus learn from our findings? One of the most important motives for closet indexing is managers' lack of additional investment ideas, and fear of losing revenue by revealing that they are not as skilled as other fund managers. These funds should therefore not be forced to increase activity, but rather update their investor information and reduce the fees they charge. Regulatory authorities should be cautious about forcing potential closet index funds to become more active. However, this is not that simple in practice because a regulator cannot stop managers from increasing activity and thus justifying their high fees. If managers, who are not adequately skilled, choose to increase their activity, poor results inevitably lead to lower assets under management. In the most severe cases, these funds will be forced out of business. However, this process is slow and uninformed investors will sustain major losses on the way.

In sum, a regulator should not force funds to become more active. Instead, their role should be

to identify closet index funds. If funds want to maintain a low degree of activity, this should be reflected in the fee levels and information set. If the trend of investigating closet index funds continues, the hope is that this will discourage new funds from choosing such a strategy. In this context, the DNB case from Norway illustrates that legal settlements can also be obtained. Intermediaries and information providers such as Morningstar should pay special attention to the performance of closet index funds that want to become a truly active fund. 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 \(2020\)](#)). Naturally, information on active share and tracking error should also be available to investors.

Mutual funds manage 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 sees more capital being allocated away from active funds and into index funds ([Investment Company Institute \(2019\)](#)).³¹ Eventually, the flow of funds from active to passive has to stop. The cost of active funds will decrease, making active funds more attractive relative to passive funds. However, we always need to pay attention to the content of the active funds. The debate on closet indexing gained momentum after the paper by [Cremers and Petajisto \(2009\)](#) and is still ongoing. The ruling against DNB in the Norwegian Supreme Court in February 2020 is a milestone in this debate. To our knowledge, this is the first time sanctions were imposed by the courts. Only the future can tell if and how the closet index problem will be resolved.

³¹There are many reasons for this development. One is the poor historical performance of actively managed funds, another is increased competition between funds due to new technology. A final one may be what we examine in this paper, the difference between what some funds promise and what they deliver.

References

- Adams, J. C., Hayunga, D. K. and Mansi, S. (2021), ‘Scale and performance in active management are not negatively related’, *Critical Finance Review*, *forthcoming*.
- 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.
- Angrist, J. D. and Pischke, J.-S. (2008), *Mostly harmless econometrics*, Princeton university press.
- Antón, M., Cohen, R. B. and Polk, C. (2021), ‘Best ideas’, Available at <https://papers.ssrn.com/sol3/papers.cfm?abid=1364827>.
- 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.
- 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.
- Biliás, 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., Fulkerson, J. A. and Riley, CFA, T. B. (2022), ‘Active share and the predictability of the performance of separate accounts’, *Financial Analysts Journal* pp. 1–19.
- Cremers, K. M., Fulkerson, J. A. and Riley, T. B. (2021), ‘Benchmark discrepancies and mutual fund performance evaluation’, *Journal of Financial and Quantitative Analysis* pp. 1–29.
- 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.
- Doshi, H., Elkamhi, R. and Simutin, M. (2015), ‘Managerial activeness and mutual fund performance’, *The Review of Asset Pricing Studies* **5**(2), 156–184.

- 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.
- Elton, E. J. and Gruber, M. J. (2020), ‘A review of the performance measurement of long-term mutual funds’, *Financial Analysts Journal* **76**(3), 22–37.
- 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.
- 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/occasional-paper-13.pdf>, Last accessed on 2020-09-10.
- Financial Conduct Authority (2019), ‘Closet trackers’. Available at <https://www.fca.org.uk/firms/authorised-and-recognised-funds/closet-trackers>.
- Financial Supervisory Authority 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 Authority 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.
- 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.

- 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>.
- 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.
- Newey, W. K. and West, K. D. (1987), ‘A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix’, *Econometrica* **55**(3), 703–708.
URL: <http://www.jstor.org/stable/1913610>
- 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’, *Journal of Financial Economics* **138**(3), 614–634.
- Pastor, L., Stambaugh, R. F., Taylor, L. A. and Zhu, M. (2021), ‘Diseconomies of scale in active management: Robust evidence’.
- Pástor, L. and Vorsatz, M. B. (2020), ‘Mutual fund performance and flows during the covid-19 crisis’, *The Review of Asset Pricing Studies* **10**(4), 791–833.
- 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. (2021), ‘How much does size erode mutual fund performance? a regression discontinuity approach’, *Review of Finance* **25**(5), 1395–1432.
- 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 Appendix describing the DNB Norge case

On June 21, 2016, the Norwegian Consumer Council instituted legal proceedings before Oslo District Court against DNB Asset Management AS (henceforth DNB), a wholly-owned subsidiary of DNB ASA providing 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 reason to expect active management. In contrast, the funds were in fact managed very close to the index.

Oslo District Court passed its judgment on January 12, 2018, whereby the claim was rejected, and DNB was held not liable. On February 12, 2018, the Norwegian Consumer Council appealed the judgment to Borgarting Court of Appeal. The Court of Appeal's ruling was announced on May 8, 2019 and ruled in favor of the Norwegian Consumer Council in the group action.³²

In short, the Court of Appeal describes the effort of active management as (i) performing analyses to identify potential good bets; and (ii) translating this into active positions such that the fund deviates from the index to a not insignificant extent. 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 investors. DNB was sentenced to pay approximately NOK 350 mill (approx. USD 35 mill).³³

DNB appealed the case to the Norwegian Supreme Court. The appeal case started on January 21, 2020. The ruling was delivered on February 27, upholding the Court of Appeal's ruling.³⁴³⁵³⁶

³²Available at [Lovdata TOSLO-2016-105341-2](#).

³³[Lovdata LB-2018-43087](#).

³⁴[Better Finance press release \(2020\)](#).

³⁵[Norwegian Consumer Council press release](#).

³⁶[Lovdata HR-2020-475-A](#).

B Appendix describing the Data

B.1 Sample selection and raw fund data

In this section, we present details of the sample selection and construction. Next, we present information on the fund data. This section is intended to help replicate the paper's results.

Sample selection

The initial sample is constructed from lists generated in Morningstar Direct. As explained in the main part, and shown in Table 1, the treated countries are Denmark, Norway, and Sweden. These are the countries where the financial authorities have applied the most intense scrutiny. The initial sample of European countries are those where we have been unable to identify any scrutiny by the FSAs: Austria, Belgium, Finland, Poland, Portugal, and Switzerland. For each of these countries, we construct lists based on the fields *Global Broad Category Group*, *Investment Area*, *Firm Country*, and *Base 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. Moreover, 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 in the same geographic area as they invest in. 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 funds to be alive one year before and after the event. Thus we exclude all funds that have an inception date after January 31, 2013 or an obsolete date before December 31, 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 lack 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 $\leq 50\%$	122	47	75
Active share $\leq 40\%$	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. As the two data providers categorize 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 final requirement is that the funds have sufficient data before and after the event window, of January 2014 until December 2015. To draw meaningful inference, we require that each fund has data starting, at the latest, one year prior to 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 the Morningstar and Lipper databases 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 returns data is in the local currency, while assets under management are in USD to establish a common currency for comparison across countries. The Monthly Return includes management, administrative, and other costs that are deducted from the NAV, and gross returns are returns before fees. Thus, we use these two variables to compute the expense ratios in accordance with Morningstar Direct definition.

For the portfolio data, we use both the Morningstar and Lipper databases. However, we find that some of the other European countries' funds lack 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 evaluate the fund portfolios and performance against. 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 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. Moreover, 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 lack 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%.

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 actively managed domestic equity mutual funds in Europe. Values are means over a two-year window before the event start. Gross alpha, expense ratio, and net alpha are annualized. Competition is defined as $1 - \text{Herfindahl-Hirschman index (HHI)}$, and displayed values are country-means for the Scandinavian countries and the rest of Europe, respectively. Scandinavian funds are labeled treated and the rest of Europe control. Inference on differences between treatment and control funds is based on [Newey and West \(1987\)](#) standard errors. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Sample	Full sample	Treated	Control	Difference
Number of funds	353	156	197	
Active share (%)	55.0	50.9	59.0	−8.1***
Gross alpha (%)	0.86	0.62	1.06	−0.44
Expense ratio (%)	1.34	1.37	1.31	0.06
Net alpha (%)	−0.50	−0.78	−0.27	−0.52
AUM (million USD)	240	336	165	171**
Fund age (years)	12.0	13.0	10.7	2.3
Competition ($1 - \text{HHI}$)	0.82	0.84	0.81	0.03

B.4 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.

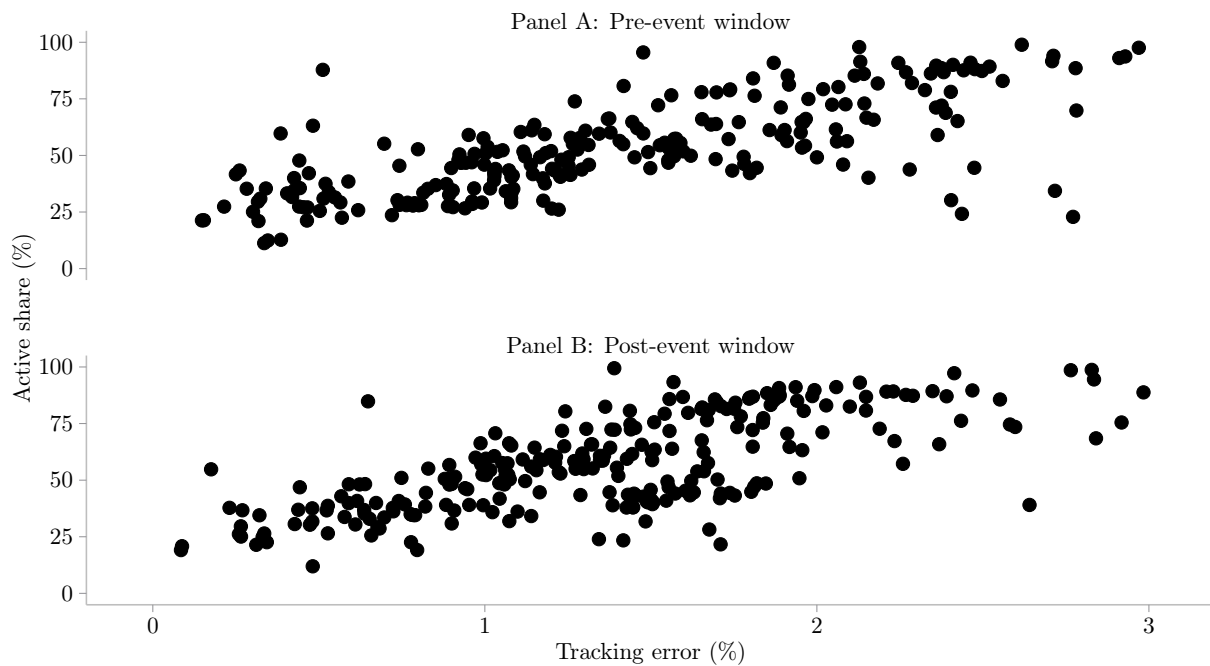
Name	Definition
Active share	Percentage of a fund's portfolio holdings that differ from its benchmark index holdings.
Expense ratio	Monthly expense ratio.
Gross alpha	Difference between the fund gross return and its benchmark return.
Net alpha	Difference between the fund net return and its benchmark return.
Factor-adjusted alpha	Three-factor alpha (percentage per month) with country-specific factors.
AUM	Total assets under management in USDm for all share classes.
Fund age	Number of years since the fund's launch date.
Industry competition	$1 - \text{Herfindah-Hirschman index (HHI)}$. HHI is the fund family-level industry concentration in the country and defined as $\text{HHI} = \sum_{i=1}^N (w_i)^2$, where w_i is the AUM-based weight of fund family i in the country.

B.5 Relationship 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 these two measures are highly correlated for the domestic funds in our sample. This confirms the findings from [ESMA \(2020\)](#).

Figure B.1.
Tracking error and active share

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



C Appendix with additional analysis for impact on active share

C.1 Testing for pre-trends

Table C.1.
Pre-trend active share

This table reports estimated slope coefficients ($\text{Post} \times t$) from panel regressions testing the presence of pre-trends in **active share** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All regressions include fund and month fixed effects, and column 2 additionally controls for fund age, size, and fees. Standard errors clustered by fund and month are in parentheses. There are 1,864 and 2,875 fund-month observations in the 40% and 50% panels, respectively. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Post \times t	
	(1)	(2)
Active share $\leq 40\%$	−0.13* (0.07)	−0.14* (0.08)
Active share $\leq 50\%$	−0.02 (0.06)	−0.03 (0.06)
Controls		×

C.2 Impact on active share based on fund characteristics

Table C.2.
Heterogeneity in the effect of intervention on active share

This table reports heterogeneous effects of policy interventions on **active share** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. Funds are classified as "high" and "low" based on the median value in their respective country. The high (low) column reports difference-in-differences estimates using a high (low) sub-sample of treated and control funds. The difference column contains estimates from a tripe difference-in-differences model. [Newey and West \(1987\)](#) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Active share $\leq 40\%$		
Fund characteristic	High	Low	Difference
Size	8.64** (3.34)	10.04*** (3.47)	-1.40 (4.86)
Age	12.41*** (2.45)	5.65 (4.01)	6.76 (4.79)
Fee	10.77*** (3.03)	8.11** (3.16)	2.66 (4.36)
Performance	9.92*** (3.42)	9.08*** (3.13)	0.85 (4.61)
Fund sample	Active share $\leq 50\%$		
Fund characteristic	High	Low	Difference
Size	7.08*** (2.65)	3.54 (2.98)	3.54 (3.94)
Age	9.26*** (2.49)	2.03 (3.10)	7.23* (3.99)
Fee	3.12 (2.60)	8.31*** (2.61)	-5.18 (3.63)
Performance	6.28** (2.52)	5.01* (2.65)	1.27 (3.64)

C.3 Impact on active share using fixed effects

Table C.3.
Active share-scrutiny relationship

This table reports estimated slope coefficients ($\text{Post} \times \text{Scrutiny}$) from panel regressions testing the effects of policy interventions on **active share** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All regressions include fund and month fixed effects, and columns 3 and 4 additionally control for fund age, size, and fees. Standard errors (in parentheses) clustered by fund are reported in columns 1 and 3 and independently by fund and month in columns 2 and 4. There are 3,711 and 5,582 fund-month observations in the 40% and 50% panels, respectively. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Post \times Scrutiny			
	(1)	(2)	(3)	(4)
Active share $\leq 40\%$	9.27*** (2.31)	9.27*** (2.31)	8.24*** (2.22)	8.24*** (2.22)
Active share $\leq 50\%$	6.50*** (1.94)	6.50*** (1.95)	5.97*** (1.87)	5.97*** (1.88)
Controls			×	×
Fund cluster	×	×	×	×
Month cluster		×		×
Adj. R ² [40% 50%]	[0.57 0.66]		[0.59 0.66]	

D Appendix with additional analysis for impact on fee

D.1 Testing for pre-trends

Table D.1.
Pre-trend fees

This table reports estimated slope coefficients ($\text{Post} \times t$) from panel regressions testing the presence of pre-trends in **fees** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All regressions include fund and month fixed effects, and column 2 additionally controls for fund age and size. Reported coefficients are annualized from fund-month level observations. Standard errors clustered by fund and month are in parentheses. There are 1,861 and 2,879 fund-month observations in the 40% and 50% panels, respectively. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Post \times t	
	(1)	(2)
Active share $\leq 40\%$	−0.00 (0.00)	−0.00 (0.00)
Active share $\leq 50\%$	−0.00 (0.00)	−0.00* (0.00)
Controls		×

D.2 Impact on fee based on fund characteristics

Table D.2.
Heterogeneity in the effect of intervention on fees

This table reports heterogeneous effects of policy interventions on **fees** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. Funds are classified as "high" and "low" based on the median value in their respective country. The high (low) column reports difference-in-differences estimates using a high (low) sub-sample of treated and control funds. The difference column contains estimates from a triple difference-in-differences model. [Newey and West \(1987\)](#) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Active share $\leq 40\%$		
Fund characteristic	High	Low	Difference
Size	0.03 (0.05)	-0.19* (0.11)	0.21* (0.12)
Age	-0.14 (0.09)	0.02 (0.03)	-0.16 (0.11)
Fee	-0.19* (0.11)	0.06* (0.03)	-0.25** (0.11)
Performance	-0.09 (0.08)	0.01 (0.05)	-0.10 (0.10)
Fund sample	Active share $\leq 50\%$		
Fund characteristic	High	Low	Difference
Size	0.04 (0.04)	-0.03 (0.07)	0.07 (0.08)
Age	-0.09 (0.06)	0.11** (0.05)	-0.20*** (0.07)
Fee	-0.04 (0.07)	0.08** (0.03)	-0.11 (0.08)
Performance	0.00 (0.07)	0.04 (0.05)	-0.05 (0.08)

D.3 Impact on fee using fixed-effects

Table D.3.
Fee-scrutiny relationship

This table reports estimated slope coefficients ($\text{Post} \times \text{Scrutiny}$) from panel regressions testing the effects of policy interventions on **fees** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All regressions include fund and month fixed effects, and columns 3 and 4 additionally control for fund age, and size. Reported coefficients are annualized from fund-month level observations. Standard errors (in parentheses) clustered by fund are reported in columns 1 and 3 and independently by fund and month in columns 2 and 4. There are 3,734 and 5,761 fund-month observations in the 40% and 50% panels, respectively. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Post \times Scrutiny			
	(1)	(2)	(3)	(4)
Active share $\leq 40\%$	−0.05 (0.05)	−0.05 (0.05)	−0.06 (0.05)	−0.06 (0.05)
Active share $\leq 50\%$	0.01 (0.04)	0.01 (0.04)	0.01 (0.04)	0.01 (0.04)
Controls			×	×
Fund cluster	×	×	×	×
Month cluster		×		×
Adj. R ² [40% 50%]	[0.92 0.92]		[0.92 0.92]	

E Additional analyses regarding impact on alpha

E.1 Testing for pre-trends

Table E.1.
Pre-trend alpha

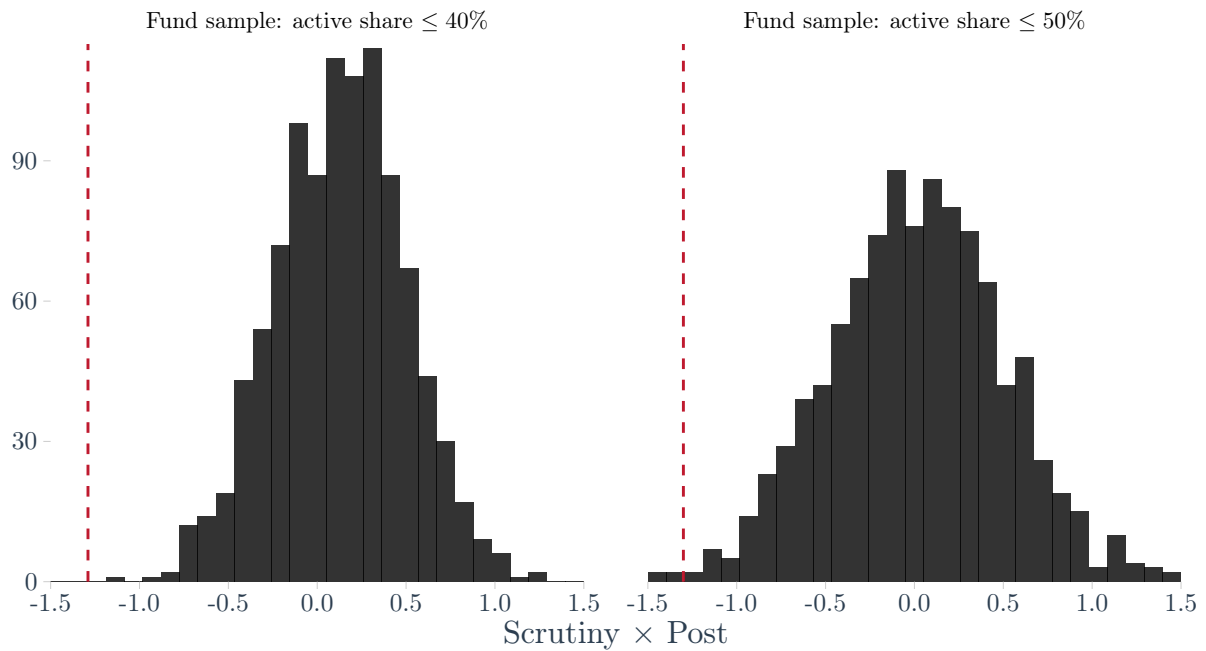
This table reports estimated slope coefficients ($\text{Post} \times t$) from panel regressions testing the presence of pre-trends in **alpha** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All regressions include fund and month fixed effects, and column 2 additionally controls for fund age, size, and fees. Reported coefficients are annualized from fund-month level observations. Standard errors clustered by fund and month are in parentheses. There are 1,859 and 2,876 fund-month observations in the 40% and 50% panels, respectively. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Post \times t	
	(1)	(2)
Active share $\leq 40\%$	−0.04 (0.16)	−0.04 (0.17)
Active share $\leq 50\%$	0.04 (0.18)	0.05 (0.20)
Controls		×

E.2 Placebo tests

Figure E.1.
Random assignment to treatment and control group

This figure presents histograms of the estimated coefficients of a falsification test for the difference-in-differences model for fund alpha, equivalent to the one presented in column 2 in Table 5. Fund sample denotes the active share cutoff limit used 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 that in the original sample (see Table 2). The model is then re-estimated using the randomly assigned treatment variable. The reported coefficients are for the interaction $\text{Scrutiny} \times \text{Post}$. The dashed vertical lines mark the corresponding coefficient estimates from Table 5 column 2.



E.3 Impact on alpha based on fund characteristics

Table E.2.
Heterogeneity in the effect of intervention on alpha

This table reports heterogeneous effects of policy interventions on **alpha** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. Funds are classified as "high" and "low" based on the median value in their respective country. The high (low) column reports difference-in-differences estimates using a high (low) sub-sample of treated and control funds. The difference column contains estimates from a tripe difference-in-differences model. [Newey and West \(1987\)](#) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Active share $\leq 40\%$		
Fund characteristic	High	Low	Difference
Size	-1.64** (0.81)	-0.95 (0.60)	-0.70 (1.02)
Age	-1.62** (0.75)	-0.90 (0.69)	-0.71 (1.09)
Fee	-0.93 (0.64)	-1.53* (0.88)	0.60 (1.11)
Performance	-1.03* (0.59)	-1.00 (0.63)	-0.04 (0.98)
Fund sample	Active share $\leq 50\%$		
Fund characteristic	High	Low	Difference
Size	-0.80 (0.72)	-1.84*** (0.67)	1.04 (0.98)
Age	-1.13 (0.69)	-1.49** (0.64)	0.36 (0.95)
Fee	-1.31** (0.57)	-1.22* (0.73)	-0.08 (0.96)
Performance	-1.27** (0.57)	-1.25** (0.61)	-0.01 (0.78)

E.4 Value creation alternatives

Table E.3.

Effect of intervention on alternative measures of value creation

This table reports the consequences of policy interventions on **net alpha** and **factor-adjusted alpha** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. Estimation is based on two-year pre- and four-year post-event averages. The difference column presents post-event differences between treated and control funds. The difference-in-differences (DiD) follows a classic setup, see [Angrist and Pischke \(2008\)](#) for details. Reported coefficients are annualized from fund-month level observations. [Newey and West \(1987\)](#) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Net alpha		Factor-adjusted alpha	
	Difference	DiD	Difference	DiD
Active share $\leq 40\%$	−0.73** (0.31)	−1.19** (0.54)	−0.41 (0.39)	−1.07** (0.48)
Active share $\leq 50\%$	−1.14*** (0.31)	−1.27*** (0.47)	−0.11 (0.38)	−0.64 (0.42)

Table E.4.
Value creation-scrutiny relationship

This table reports estimated slope coefficients (Post \times Scrutiny) from panel regressions testing the effects of policy interventions on **net alpha** (panel A) and **factor-adjusted alpha** (panel B) for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All regressions include fund and month fixed effects, and columns 3 and 4 additionally controls for fund age, size, and fees. Reported coefficients are annualized from fund-month level observations. Standard errors (in parentheses) clustered by fund are reported in columns 1 and 3 and independently by fund and month in columns 2 and 4. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

	Panel A: Net alpha			
Fund sample	(1)	(2)	(3)	(4)
Active share $\leq 40\%$	−1.09** (0.48)	−1.09 (1.13)	−0.94* (0.51)	−0.94 (1.22)
Active share $\leq 50\%$	−1.11** (0.43)	−1.11 (1.29)	−1.09** (0.45)	−1.09 (1.35)
Controls			×	×
Fund cluster	×	×	×	×
Month cluster		×		×
Adj. R ² [40% 50%]	[0.10 0.13]		[0.10 0.13]	
	Panel B: Factor-adjusted alpha			
Fund sample	(1)	(2)	(3)	(4)
Active share $\leq 40\%$	−0.95* (0.49)	−0.95 (1.08)	−0.69 (0.51)	−0.69 (1.14)
Active share $\leq 50\%$	−0.56 (0.42)	−0.56 (1.03)	−0.46 (0.43)	−0.46 (1.06)
Controls			×	×
Fund cluster	×	×	×	×
Month cluster		×		×
Adj. R ² [40% 50%]	[0.09 0.10]		[0.09 0.10]	

F Additional analyses regarding signal jamming

Table F.1.
Portfolio concentration-scrutiny relationship

This table reports estimated slope coefficients ($\text{Post} \times \text{Scrutiny}$) from panel regressions testing the effects of policy interventions on **portfolio concentration** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All regressions include fund and month fixed effects, and columns 3 and 4 additionally controls for fund age, size, and fees. Standard errors (in parentheses) clustered by fund are reported in columns 1 and 3 and independently by fund and month in columns 2 and 4. There are 3,709 and 5,577 fund-month observations in the 40% and 50% panels, respectively. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Post \times Scrutiny			
	(1)	(2)	(3)	(4)
Active share $\leq 40\%$	0.61*** (0.21)	0.61*** (0.22)	0.50** (0.22)	0.50** (0.23)
Active share $\leq 50\%$	0.34 (0.21)	0.34 (0.22)	0.29 (0.21)	0.29 (0.22)
Controls			\times	\times
Fund cluster	\times	\times	\times	\times
Month cluster		\times		\times
Adj. R ² [40% 50%]	[0.76 0.82]		[0.77 0.82]	

G True active Scandinavian funds 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 previously.

In Table G.1, we present the same summary statistics as we did in Table 2, but the control group is now truly active funds. There are three main differences between the samples. Firstly, active share is lower for the closet index funds. Secondly, even if not by 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 may often be old funds with uninformed investors.

Table G.1.
Summary statistics: Scandinavian funds

This table presents summary statistics for the sample of actively managed domestic equity mutual funds in Scandinavia (Norway, Sweden, and Denmark). Scandinavian non-closet indexers are assigned to the control group. Values are means over a two-year window before the event start. Gross alpha, expense ratio, and net alpha are annualized. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. Inference on differences between treatment and control funds is based on Newey and West (1987) standard errors. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample Group	Active share $\leq 40\%$			Active share $\leq 50\%$		
	Treated	Control	Difference	Treated	Control	Difference
Number of funds	46	104		75	75	
Active share (%)	32.4	61.4	-29.0***	38.0	67.0	-29.0***
Gross alpha (%)	1.04	0.77	0.27	1.24	0.45	0.79
Expense ratio (%)	1.15	1.44	-0.29***	1.24	1.46	-0.22***
Net alpha (%)	-0.14	-0.73	0.59	-0.05	-1.05	1.00
AUM (million USD)	476	283	193*	379	306	73
Fund age (years)	16.2	12.0	4.2***	15.0	11.6	3.4***

In Table G.2 and G.3 we present how scrutiny influence closet index funds relative to truly active funds.

Table G.2.
Robustness: effect of intervention on main outcomes

This table reports the effects of policy interventions on **active share**, **fees** and **alpha** for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. The control group comprises Scandinavian non-closet indexers. Values for fees and alpha are annualized. Estimation is based on two-year pre- and post-event averages (for alpha four-year post event). The difference column presents post-event differences between treated and control funds. The difference-in-differences (DiD) follows a classic setup, see [Angrist and Pischke \(2008\)](#) for details. For fee and alpha, reported coefficients are annualized from fund-month level observations. [Newey and West \(1987\)](#) standard errors in parentheses. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

Fund sample	Active share		Fees		Alpha	
	Difference	DiD	Difference	DiD	Difference	DiD
Active share $\leq 40\%$	-30.32*** (2.48)	3.16* (1.86)	-0.41*** (0.07)	-0.09* (0.05)	-1.62*** (0.50)	-2.21*** (0.84)
Active share $\leq 50\%$	-26.96*** (2.09)	1.30 (1.72)	-0.27*** (0.06)	-0.05 (0.04)	-1.46*** (0.50)	-2.25*** (0.80)

Table G.3.
Robustness: main outcomes-scrutiny relationship

This table reports estimated slope coefficients (Post \times Scrutiny) from panel regressions testing the effects of policy interventions on **active share** (panel A), **fees** (panel B), and **alpha** (panel C) for closet index funds. Fund sample denotes the active share cutoff limit used to classify funds as closet indexers. All regressions include fund and month fixed effects, and columns 3 and 4 additionally controls for fund age, size, and fees (not in panel B). For fee and alpha, reported coefficients are annualized from fund-month level observations. Standard errors (in parentheses) clustered by fund are reported in columns 1 and 3 and independently by fund and month in columns 2 and 4. Asterisks denote statistical significance: *** $p < 1\%$, ** $p < 5\%$, * $p < 10\%$.

	Panel A: Active share			
Fund sample	(1)	(2)	(3)	(4)
Active share $\leq 40\%$	3.86* (2.02)	3.86* (1.99)	3.39 (2.09)	3.39 (2.05)
Active share $\leq 50\%$	1.27 (1.85)	1.27 (1.83)	0.78 (1.93)	0.78 (1.89)
Controls			×	×
Fund cluster	×	×	×	×
Month cluster		×		×
Adj. R ² [40% 50%]	[0.78 0.77]		[0.77 0.77]	
	Panel B: Fees			
Fund sample	(1)	(2)	(3)	(4)
Active share $\leq 40\%$	−0.07 (0.06)	−0.07 (0.05)	−0.04 (0.05)	−0.04 (0.05)
Active share $\leq 50\%$	−0.03 (0.04)	−0.03 (0.04)	−0.02 (0.04)	−0.02 (0.04)
Controls			×	×
Fund cluster	×	×	×	×
Month cluster		×		×
Adj. R ² [40% 50%]	[0.83 0.83]		[0.85 0.86]	
	Panel C: Alpha			
Fund sample	(1)	(2)	(3)	(4)
Active share $\leq 40\%$	−1.96** (0.82)	−1.96 (1.31)	−1.82** (0.83)	−1.82 (1.31)
Active share $\leq 50\%$	−2.08** (0.81)	−2.08* (1.24)	−2.09** (0.82)	−2.09* (1.25)
Controls			×	×
Fund cluster	×	×	×	×
Month cluster		×		×
Adj. R ² [40% 50%]	[0.09 0.11]		[0.09 0.11]	