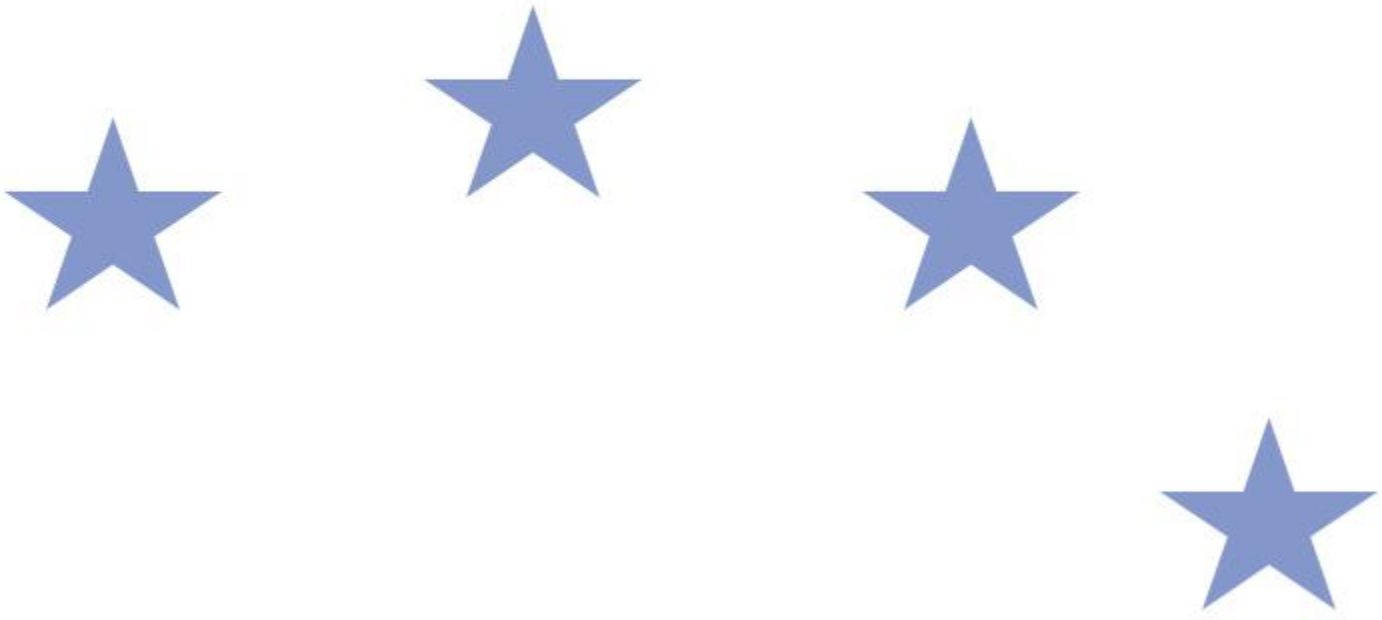
esma European Securities and Markets Authority

ESMA Working Paper
No. 2, 2020

Closet indexing indicators and investor outcomes

Lorenzo Danieli, Alexander Harris and Giorgia Pichini



September 2020 | ESMA50-165-1292

ESMA Working Paper, No. 2, 2020

Authors: Lorenzo Danieli, Alexander Harris and Giorgia Pichini
Authorisation: This Working Paper has been approved for publication by the Selection Committee and reviewed by the Scientific Committee of ESMA.

European Securities and Markets Authority (ESMA)
Risk Analysis and Economics
201-203 rue de Bercy
FR–75012 Paris
risk.analysis@esma.europa.eu

Closet indexing indicators and investor outcomes♦

Lorenzo Danieli, Alexander Harris, Giorgia Pichini ♣

Abstract

This paper examines how different measures of closet indexing relate to costs and performance among EU-domiciled equity funds with active mandates. To capture different hallmarks of closet indexing we study a range of metrics, conducting panel regressions using annual fund-level data for 2010-2018. In addition to the portfolio-based Active Share we use three return-based metrics: Tracking Error, R^2 and Beta. We develop a combined indicator to study the joint effects of the return-based metrics and verify that it predicts Active Share. Turning to our main results, potential closet indexing is associated with lower performance, including for risk-adjusted performance (alpha) net of costs. Potential closet indexing is associated with slightly lower total costs, though far above those for passive funds. Closet indexers therefore appear to pass on to consumers only a small share of the lower economic costs of benchmark-tracking compared to active management, rather than engaging in price competition.

JEL Classifications: G18, G23, G28

Keywords: Active funds, active management, closet indexing, fees, index funds, index tracking, investor protection, passive management, performance

♦ The views expressed are those of the authors and do not necessarily reflect the views of the European Securities and Markets Authority. Any errors or omissions are the responsibility of the authors. The authors would like to thank Manuela Geranio for her detailed feedback. They would also like to thank Isabella Cortesi, Damien Fennell, Massimo Ferrari, Claudia Guagliano, Steffen Kern, Tania de Renzis, Christian Winkler and participants in a workshop with NCAs held at ESMA in January 2020 for their comments and suggestions. The authors are grateful to Yakov Amihud, Martijn Cremers and Antti Petajisto for their helpful explanations via correspondence.

♣ European Securities and Markets Authority (ESMA), 201-203 rue de Bercy, 75012 Paris Cedex, France. Email contacts: lorenzo.danieli@esma.europa.eu, alexander.harris@esma.europa.eu, giorgia.pichini@esma.europa.eu

1. Introduction

Many funds are managed in reference to a benchmark index. If a fund's objective is to track an index, then it is commonly referred to as a passive fund.[1] Alternatively, if the fund's mandate allows the manager at least some discretion beyond index tracking, the fund may be referred to as an active fund.[2]

Among those active funds that use a benchmark index, the index may play a role in the management of the fund in different ways. For example, an active fund may aim to outperform its benchmark (possibly by a minimum amount and/or over a specified time period) or may use its benchmark to define its investment universe.

This paper analyses the specific situation in which asset managers claim to manage their funds in an active manner while in fact tracking or staying close to a benchmark index. Such practices – often referred to as "closet indexing" – represent a form of misconduct on the part of the asset manager, which has been criticised by supervisors and investor advocacy groups on numerous occasions in recent years.[3] An economic incentive for closet indexing is that fees for funds with an active mandate tend to be higher than fees for passive funds, while the economic costs of passive management may be less than for active management, since staying close to a benchmark requires less input from the manager. While supervisory activities are ongoing to limit such practices, this paper seeks to contribute to the understanding of closet indexing in the EU by providing evidence on how closet indexing relates to the costs and performance of EU-domiciled equity funds.

1.1 Policy context

If a fund manager does carry out closet indexing, this has consequences for investor protection.[4] An immediate concern is that, by definition, closet indexing involves misinforming prospective and current investors. Information in key disclosure documents such as the fund's Prospectus and Key Investor Information Document (KIID) may be unclear, inaccurate or incomplete.[5]

Closet indexing also raises other, related concerns, as follows.

- Investors could be making investment decisions based on an inaccurate expectation of receiving a more active fund management service than they actually will.
- Investors may be exposed to a different risk/return profile than they had envisaged.
- Investors may also be subject to higher fees than they would pay for a passive fund that explicitly tracked a given benchmark index.

Identifying the extent to which closet indexing takes place in the European collective investment industry is therefore an important question, and an ongoing one. In recent years, ESMA and

[1] No definition of active and passive funds (UCITS or AIFs) exists in EU Level 1 or Level 2 legislation. ESMA has published (non-legally-binding) Q&A on the application of the UCITS Directive (ESMA, 2019). The Q&A state that an actively managed UCITS is one where the manager has discretion over the composition of its portfolio, subject to the stated investment objectives and policy. As opposed to a passive UCITS, an active UCITS does not have an index-tracking objective although it may include or imply reference to a benchmark.

[2] To distinguish funds with an active mandate that are not closet indexers from those that are, we sometimes refer to the former as 'funds that pursue active strategies'. Unless otherwise stated, 'actively managed funds' and 'active management' include potential closet indexers.

[3] See e.g. ESMA (2016), Central Bank of Ireland (2019), Better Finance (2019a).

[4] See ESMA (2016).

[5] The KIID is a document containing key investor information, to help investors understand the nature and the risks of investing in a given fund. Commission Regulation 583/2010 ('KIID Regulation') provides a harmonised regime on the form and content of the disclosure ensuring that information about investment opportunities in the UCITS market is consistent and comparable. KIIDs must include sections on objectives and investment policy, risk and reward profile, charges, past performance and practical information.

National Competent Authorities (NCAs) have worked to identify potential closet indexers by examining metrics on fund composition and performance and by conducting follow-up detailed supervisory work on a fund-by-fund basis.[6] Throughout this work, ESMA has recognised that data-based metrics used to identify potential closet indexers among large samples of funds, while imperfect screening tools, are nonetheless a useful source of evidence to help direct supervisory focus.[7]

Closet indexing is fundamentally a concern for investors because it involves misrepresenting a fund's investment strategy. Poor performance by funds is an additional, general issue for investors. The results in this paper suggest that indicators of potential closet indexing are associated with slightly reduced performance *on average* compared to genuinely active funds. This result holds on a risk-adjusted basis net of fees, suggesting that the core investor protection concerns outlined above around closet indexing are not offset by an improved risk-return profile among potential closet indexers. Ex-ante, investors seeking an actively managed fund are worse off with a closet indexer, even when costs are taken into account.

Active management inevitably involves risks. Among funds that pursue genuinely active strategies, some may materially underperform their benchmarks. In other words, they would have received higher returns from tracking their benchmarks. However, other genuinely active funds may materially outperform their benchmarks. Active management in general offers the possibility that a fund may outperform the market. Closet indexing does not offer the same ex-ante risk profile that investors should expect from genuine active management. In particular, it does not offer scope for strongly positive alpha (i.e. returns adjusted for market risk, defined more precisely in section 1.2).

A point of contrast with the misinformation entailed by closet indexing is that persistent underperformance by genuinely active funds can typically be directly identified by publicly available information.[8] To help investors protect themselves against inadequate fund performance and excessive fees, ESMA publishes its Annual Statistical Report on Performance and Costs of Retail Investment Products in the EU.[9] The report provides extensive comparative statistics on an annual basis about gross and net returns of fund products sold in the EU, which can serve as an important point of orientation for investors. The report also presents yardsticks for the performance of funds over several time horizons.

1.2 Research focus

The primary focus of this paper is to provide evidence on how closet indexing relates to the costs and performance of EU-domiciled equity funds. To our knowledge, our study is the first to examine closet indexing and investor outcomes for EU-domiciled equity funds for all active mandates, and the first to investigate return-based measures of closet indexing across this market.

We report results both for portfolio- and return-based metrics of closet indexing and in general frame our results in terms of 'potential closet indexers', i.e. funds whose portfolio and/or performance data indicate they are more likely than other funds to carry out closet indexing. Given the economic incentives for closet indexing highlighted above, a natural hypothesis – and one which has been tested in the literature, mostly in the context of US-domiciled funds – is that

[6] For a summary of work in this area, see the ESMA Report on Trends, Risks and Vulnerabilities No. 2, 2020, pp. 37-38

[7] Kjørven (2019) highlights the benefits of developing a coordinated approach to identifying potential closet indexing at a European level.

[8] In the EU, UCITS managed against a benchmark are required to disclose past performance of the fund alongside that of the benchmark. See Article 18(1) of Commission Regulation (EU) No 583/2010 for details.

[9] See ESMA (2020).

potential closet indexers tend to charge fees similar to other active managers and much higher than (explicitly) passive managers. We find support for this hypothesis based on data for EU-domiciled funds.

As set out in more detail in the literature review in section 2, some studies have found that among samples of funds that state they are actively managed, certain measures of closet indexing are associated with lower performance. We report results for different measures of performance and find some continuity with this existing strand of the literature. Our primary measure of performance is a fund's (Jensen's) alpha, net of costs. Alpha is a commonly-used performance measure in the literature as it identifies net residual performance beyond that accounted for by benchmark exposure.[10] It therefore adjusts a fund's performance based on that of the overall market. We find that the closet indexing metrics we study are associated with lower alpha, whether or not controls are included in the regression. The same is true for unadjusted gross returns. Turning to costs, we find that potential closet indexers are associated with a slightly lower Total Expense Ratio (TER), though TER remains far higher than for a typical (explicitly) passive fund.

In short, our results suggest that potential closet indexers tend to realise returns slightly below those for other active funds. They tend to charge slightly lower fees than other active funds, though considerably higher than the average fees charged by passive funds. In other words, our results confirm the concerns of supervisors and investor advocacy groups that investors in closet indexing funds face an unjustifiably high level of costs given that they do not receive what they are promised to obtain in return, namely an actively managed fund portfolio.[11]

The paper is organised as follows. Section 2 reviews the relevant literature and introduces some definitions. Section 3 sets out the empirical strategy including regression specifications. Section 4 gives an overview of the data. Section 5 presents and discusses the results of the statistical and regression analysis, which centre on how a range of measures of potential closet indexing relate to the performance and costs of EU equity funds. Section 6 concludes.

2. Literature review

Different indicators have been proposed in the literature to help identify closet indexing, as reviewed in section 2.1 and formally defined in Box 1 and Annex 2. Section 2.2 surveys existing findings in relation to the costs and performance of potential closet indexers within the context of wider research around costs and performance of active versus passive fund management.

2.1 Methods to identify potential closet indexing

For a given fund and benchmark, the two main sources of data that can be used to try to identify closet indexing are the portfolio composition of the fund versus its benchmark and the fund's performance versus that of its benchmark. Neither source of data yields perfect identification of closet indexing, however, for two reasons.

First, the portfolio composition and returns of potential closet indexers will in general differ from those of their benchmarks to varying extents, as perfect index replication is not generally

[10] Formally, Jensen's alpha – the measure used in this paper – is the intercept in a regression of fund returns on benchmark returns. 4-factor alpha is analogously defined with respect to the Carhart model set out in Box 2.

[11] The fact that passive funds do not face costs of researching and implementing active strategies enables them to charge far lower fees than funds with active mandates. For example, the ESMA Report on Trends, Risks and Vulnerabilities No. 2, 2019 reports a large difference in average TER between active and passive UCITS funds. Among passive UCITS funds in 2018, average TER was around 0.3 pp, compared with 1.42 pp among active UCITS funds.

feasible. Consequently, even if a manager tracks an index as closely as possible, the fund's portfolio will inevitably differ to some extent from the index. Returns too will differ, not only because of the portfolio differences but also because of costs.[12] Additionally, some managers may follow a strategy of partial index replication, while retaining some degree of active management, possibly to a varying extent over time.

Second, as set out in Petajisto (2009), active funds may pursue strategies that do not simply aim to replicate an index but that nonetheless closely match the benchmark in terms of portfolio composition or returns. As a result, any metric used to identify potential closet indexers is likely to yield false positives.

Different metrics have been developed to tackle the problem of identifying potential closet indexers. The main portfolio-based measure is Active Share (AS), introduced by Cremers and Petajisto (2009). Intuitively, the calculation of AS makes use of the fact that any fund can be decomposed into a benchmark component and a residual zero-net-investment long-short portfolio. The AS is simply the size of the latter.[13] This ready interpretation of AS makes it a useful metric for studying determinants of fund performance, as it indicates the *potential* for outperformance. More precisely, non-zero AS is a necessary condition for a fund to be able to outperform its benchmark, as one portfolio can only beat another by deviating from it.

However, for the reasons discussed above – namely that closet indexers will in practice have non-zero AS, and that different active strategies may involve different values of AS – it is only a proxy for active management. As such (and as Cremers and Petajisto acknowledge), studies of active management and in particular closet indexing are likely to benefit from using other metrics alongside AS. Another limitation is that in the EU (in contrast with the US), AS is reported on a voluntary basis rather than subject to mandatory disclosure, which has the added effect of limiting data availability.[14]

A commonly-used performance-based metric to help tackle the problem of identifying potential closet indexing is Tracking Error (TE), the standard deviation of the difference in fund returns and benchmark returns over time. Alternative measures of TE are based on the residuals of regressions of fund returns on factors such as those in Fama and French (1993) and Carhart (1997).[15]

[12] For instance, as the relative market values of asset holdings change over time, portfolio composition will change, necessitating rebalancing. Management of inflows and the need to meet redemptions likewise change portfolio composition due to constraints on the timing of trades for different assets and differences in transaction costs by asset. Deviations from index portfolio composition will in turn generate deviations in fund returns, both directly and via associated transaction costs from rebalancing.

[13] AS is therefore bounded above by 100% and below by 0% for unleveraged funds.

[14] Cremers and Curtis (2016) recommend that AS be included in any disclosure regime alongside costs and performance, on the basis that this may improve transparency and investor awareness.

[15] For example, Cremers and Petajisto define TE to be the standard deviation of the residuals obtained from regressing fund returns on benchmark index returns plus a constant, where returns are adjusted for the risk-free rate. This approach is designed to remove sensitivity of the metric to persistent long-term strategies. For example, returns from a long-term fixed holding of low-beta stocks would largely be reflected in the constant (alpha) term of the regression rather than the residual. In this paper, the R^2 and beta from such a regression are included in a broad set of metrics alongside the simpler, standard measure of TE. We thereby aim to capture a range of different aspects of active management in studying the impact of closet indexing on investor outcomes.

Box 1

Tracking Error (TE) and Active Share (AS)

Two key measures capture different features of active management

The AS of a fund captures the extent to which its portfolio differs from its benchmark index:

$$Active\ Share\ = \frac{1}{2}\sum_{i=1}^{N}\left|w_{fund,i} - w_{index,i}\right|$$

where $w_{fund,i}$ and $w_{index,i}$ are the portfolio weights of asset i in the fund and in the index respectively, and the sum is taken over all assets.[16] The absolute value of differences in these weights is taken so that over- and under-weightings in the fund relative to the benchmark count equally, while the normalisation factor of one half ensures that a fund that has no overlap with an index has an AS of 100%.

TE is defined as the standard deviation of the difference between a fund return and its benchmark index return.[17]

$$Tracking\ Error\ = Stdev\left(R_{fund,t} - R_{index,t}\right)$$

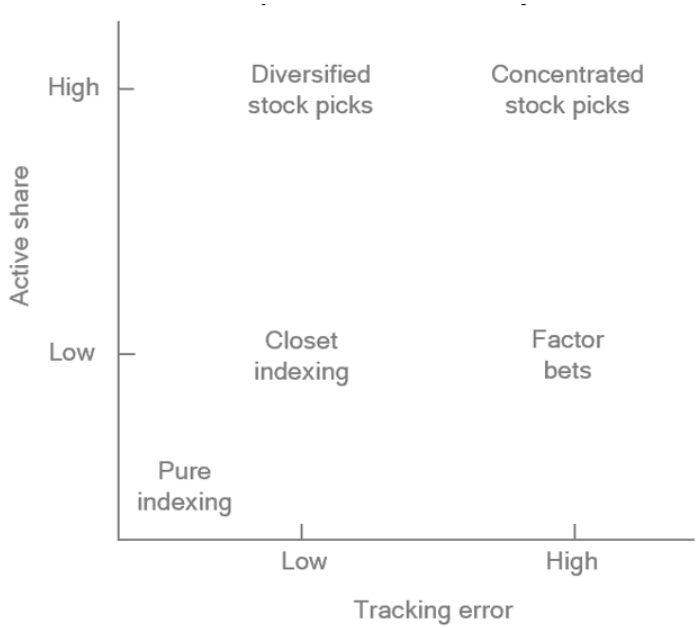
where $R_{fund,t}$ and $R_{index,t}$ are respectively the returns of a fund and its benchmark index in period $t \in \{1, \ldots, \tau\}$, and $Stdev(.)$ is the standard deviation taken over this period. TE reflects the extent to which a fund manager takes on risk from the active component of their portfolio that is not diversified across a given market. For example, if an active manager has a relatively high exposure to one industry or economic sector within a market ('factor bets', in the terminology of Chart 1), this is likely to lead to a greater TE than if the manager selects stocks with sectoral exposures reflecting those of the market as whole ('diversified stock picks'). TE is expressed in the same units as the returns of the fund and the index, typically percentage points. Because fund returns expressed as percentages are typically single digit or low double digit in absolute magnitude, TE typically ranges from near-zero (e.g. for a passive fund) to approximately 10% or 20% (Table 1 and Chart 2).

Cremers and Petajisto (2009) note that TE is sensitive to strategic decisions around factors such as momentum or value, which involve a fund manager taking correlated active positions. In contrast, AS weights all active positions equally regardless of the extent to which they are diversified. For this reason, it is likely to be more suitable as a proxy for undiversified stock picking. To the extent that TE and AS reflect these two fundamental approaches to active fund management, they are complementary. Chart 1 illustrates this analytical relationship.

[16] As in Cremers and Petajisto (2009), we restrict attention to equity funds and so the sum is taken over all equities held by funds or constituting part of a benchmark index.

[17] See e.g. Grinold and Kahn (1999).

Chart 1
Different styles of Active Management and their schematic relation to Active Share and Tracking Error



Source: Reproduced from Cremers and Petajisto (2009).

Another performance-based approach is Style-Shifting Activity (SSA), developed by Hermann et al (2016). SSA is based on the 4-factor Carhart model and measures the extent to which a fund changes its quarterly aggregate exposure to the Carhart factors of market exposure, value, size and momentum.[18] Intuitively, the 'factor bets' style of active management displayed in Chart 1 involve changes in exposure to such factors in response to a changing investment environment. Closet indexers, in contrast, might be expected to have low SSA values as closet indexing by definition involves benchmark replication.

Other performance-based metrics arise from regressing fund returns on factors. In this paper, following Fong (2016) we define a fund's R^2 to be the coefficient of determination from the Capital Asset Pricing Model (CAPM), i.e. a time series regression of fund returns on benchmark returns plus a constant. A related measure is beta, the coefficient on the benchmark returns in the same regression. By allowing for a constant in the regression, R^2 and beta are sensitive to shorter term strategic decisions as opposed to persistent long-term strategies. For example, returns from a long-term fixed holding of low-beta stocks would largely be reflected in the constant (alpha) term of the regression rather than in the residual (from which R^2 is calculated) or the beta coefficient.[19] Variations on this approach include analogously defining R^2 and beta with respect to the 4-factor Carhart model, thereby also controlling for factor exposure when calculating fund-benchmark return deviations. (Box 2 provides more information on the 4-factor model.) For example, Amihud and Goyenko (2013) use 4-factor beta to study the impact of closet indexing on costs and performance in the US market.

In a study published by the Autorité des Marchés Financiers (AMF), Demartini and Mosson (2018) calculate SSA, 4-factor TE and 4-factor R^2 for a sample of nearly 800 French funds invested in European equities. They show that the three metrics are complementary in that they exhibit covariance within the sample.

[18] For a formal definition of SSA, see Annex 2.

[19] The alpha term is the primary measure of fund performance we use in this paper.

Box 2
4-factor model (Carhart, 1997)
Four factors help explain fund performance

The 4-factor model constructed by Carhart (1997) builds on seminal studies by Fama and French (1993) and Jegadeesh and Titman (1993). One interpretation offered by the author is that the model attributes performance based on four elementary strategies: high versus low beta stocks, large versus small market capitalisation stocks, value versus growth stocks, and one-year return momentum versus contrarian stocks. The model is as follows:

$$r_{i,t} = \alpha_i + b_i RMRF_t + s_i SMB_t + h_i HML_t + p_i PR1YR_t + e_{i,t}$$

$t = 1, 2, \ldots, T$

where $r_{i,t}$ is the return on a portfolio in excess of the risk-free rate; $RMRF_t$ is the excess return on a value-weighted aggregate market proxy (as used in the CAPM); and SMB_t, HML_t, and $PR1YR_t$ are excess returns on value weighted, zero-investment, factor-mirroring portfolios for size, book-to-market equity, and one-year momentum in stock returns. For example, SMB_t gives the excess return on a portfolio which is long stocks among the bottom 30% by market capitalisation and short stocks in the top 30%; the portfolio is constructed to control for book-to-market equity, as described in Fama and French (1993). Each coefficient (b_i, s_i, h_i, p_i) represents the contribution of each strategy to fund performance.

The portfolios $RMRF_t$, SMB_t and HML_t may be thought of as risk factors, with the coefficients b_i, s_i and h_i then interpreted as risk premia. Carhart shows that the momentum factor $PR1YR_t$ is positively associated with excess returns despite not being a source of risk.[20]

2.2 Performance and costs

A theoretical framework for understanding the average returns of active funds versus benchmarks is found in Sharpe (1991).[21] Sharpe argues that the aggregated holdings of equity funds in a given market should equal the market as a whole. The benchmark index for that market represents these same aggregate holdings. The weighted average gross performance (i.e. fund return before costs) of actively managed funds should therefore equal that of the benchmark. Taking costs into account, actively managed funds should on average therefore underperform their benchmarks. Several empirical studies have supported this hypothesis. For example, Standard and Poor's (2019) conduct high-level comparisons of equity fund performance against regional indices, finding that in 2019, 71% of European active equity funds underperformed the S&P Europe 250 index. Over 3-year and 5-year horizons the figures were 79% and 78% respectively.[22] Nonetheless, the empirical evidence varies across years and regions.[23] An implication for the present context is that on average, if Sharpe's hypothesis is correct and assuming that closet indexers closely replicate their benchmarks, gross returns for

[20] Another result in Carhart (1997) is that transaction costs are sufficient to prevent investors from profitably pursuing high one-year return momentum as a strategy. Funds that happen to have a high one-year return momentum tend to realise slightly greater excess returns than other funds, controlling for the other factors in the model, but those that seek to acquire high-momentum stocks see these gains cancelled out by the associated expenses arising from the necessary trades. The existence of a momentum premium $p > 0$ may therefore be consistent with an equilibrium framework, even though momentum is not a risk factor.

[21] For the avoidance of doubt: Sharpe (1991) covers a separate topic from the 'Sharpe ratio', developed by the same author. The latter is an influential and widely-cited risk-adjusted performance measure.

[22] Standard and Poor's Index Versus Active (SPIVA) scorecards are published regularly, analysing the performance of active funds in different regions. S&P (2019) is a recent example, comparing average performance of EUR-denominated funds in Europe against a single regional index, the S&P Europe 250. In 2019, 71% of these funds underperformed the index.

[23] A recent example is Morningstar Research (2019), based on a sample of US funds representing 64% of the overall market. The study compares performance of active funds against that of passive funds, as opposed to the performance of benchmark indices, thereby taking into account the costs involved in passive management when assessing relative performance. According to the results, 48% of active U.S. stock funds survived and outperformed their average passive peer for the period July 2018 to June 2019, suggesting that the hypothesis of expected underperformance by active funds relative to passive does not always hold. In comparison, the study found that only 37% of the sample of active funds outperformed their passive peers in the previous 12-month period.

closet indexers in a given market should approximately equal those of funds that pursue active strategies.

Importantly however, the framework in Sharpe (1991) rests on certain assumptions that may not hold in reality, or only hold in approximation, depending on the market in question. The assumption that active funds' aggregate holdings are equal to the market as a whole may not hold if there exist sizeable direct equity holdings that differ in aggregate from the overall portfolio held by investment funds in the market.[24] Another reason is that equity indices may not accurately represent the market as a whole, due to time lags in being updated to reflect (continually changing) equity valuations. By the same token, the aggregate portfolio of passive funds in a given market may differ from that of active funds. In such cases, there is scope for closet indexers to outperform or underperform (either in absolute or benchmark-relative terms) funds pursuing active strategies.

Empirical studies in the academic literature on how closet indexing relates to performance have focused largely on US equity funds. To our knowledge, our study is the first to examine closet indexing and investor outcomes for EU-domiciled equity funds for all active mandates, and the first to investigate return-based measures of closet indexing across this market. Previous work by Morningstar Research (2016) has focused on Europe-domiciled funds investing in large cap equities specifically, using data from 2006 to 2015. The study finds that large cap equity funds in the top quintile of AS tended to enjoy higher average benchmark-adjusted returns than other funds.[25] A recent study that examines active management among funds internationally is that of Cremers et al (2016), who find that explicit indexing and closet indexing are associated with countries' regulatory and financial market environments.

A common pattern is that among US equity funds with an active mandate, greater activeness is associated with higher returns. Cremers and Petajisto (2009), including TE in their controls, find that US equity funds with the highest AS outperform their benchmarks, and that this outperformance is persistent. In contrast, funds with the lowest AS that nonetheless have an active mandate underperform their benchmarks. Petajisto (2013) reports similar findings, both for excess returns relative to benchmarks only and those adjusted for benchmarks and factors. Among another samples of US equity funds, Amihud and Goyenko (2013) find that activeness as inferred from 4-factor R^2 predicts excess (benchmark- and factor-adjusted) returns under different specifications. Cremers and Pareek (2015) find only qualified support for the claim that AS is associated with higher performance, however. In their study, AS is associated with higher performance only among funds whose strategies are 'patient', in the sense of a holding duration of over 2 years.[26] Finally, in contrast with the aforementioned studies, Frazzini et al (2016) find that although higher-AS US funds performed better than their lower-AS counterparts after controlling for benchmarks, the difference was not significant. The lack of a fully consistent picture in the literature of the relationship between AS and performance is one reason to study other metrics of potential closet indexing alongside AS, and to add to the body of evidence on the relationship between AS and costs by examining the EU market.

Turning to the issue of costs, Cremers et al (2016) find that actively managed funds are more active and charge lower fees when they face more competitive pressure from low-cost (explicitly) passive funds. Although competition is not the direct focus of our paper, this result is relevant to interpreting our results in suggesting that active fund managers are responsive to

[24] Fama and French (2010) find that the assumption does hold in US equity markets, i.e. the aggregate portfolio of active funds closely matches the market as whole.

[25] Better Finance (2019b) includes a regression of Jensen's alpha on TE, benchmark returns and costs for a sample of funds in Belgium, France and Luxembourg. The study funds a positive relationship between TE and alpha.

[26] This finding suggests that measures of activeness such as SSA may neglect a relevant source of alpha among funds that pursue active strategies.

market conditions when setting prices. We find that metrics associated with closet indexing are associated with slightly lower fees, even though closet indexers are likely to bear significantly lower economic costs than managers that pursue active strategies. The relatively high level of fees among active funds compared to their passive counterparts is an economic incentive to carry out closet indexing. However, closet indexers may find it optimal to set prices slightly below the average for active managers to attract inflows. Our results are in line with Amihud and Goyenko (2013), who find that among funds with active mandates, activeness as inferred from 4-factor R^2 is associated with slightly higher fees.

Box 3
Types of costs faced by investors
TER captures many types of costs

The measure of costs we employ is Total Expense Ratio (TER). Expressed as a percentage of the fund's total Net Asset Value (NAV), TER includes ongoing costs such as:

- management fees;
- operating expenses;
- legal costs;
- audit costs; and
- tax.

TER does not include certain one-off costs, such as entry and exit fees charged to investors. ESMA's analytical work on the costs and performance of long-term retail investment products (see e.g. ESMA, 2020) includes the estimation of proxies for entry and exit fees.

Cost classification varies across EU countries. Fees (in relation to management, distribution and administration) may be classified differently under national legislation, and market practices may differ by country. This can have an impact on the costs captured by TER. For example, in Italy management fees include distribution fees. In France, investment management commissions are aggregated with a number of administrative costs. In Belgium, distributor remuneration may be partly contained within the management fee of a UCITS.

---

3. Empirical approach

As set out in section 2.1, various measures of closet indexing have been developed in the literature to capture different ways in which closet indexing may be identifiable from the data. Reflecting this, ESMA's previous work on the topic has involved different data-based criteria for identifying potential closet indexing.

Our empirical strategy in this paper is based on several complementary measures of potential closet indexing. We use four of the data-based metrics discussed in section 2.1 – AS, TE, R^2 and beta – and investigate how they relate to costs and performance. Unlike in a supervisory context, our aim is not to identify precisely which funds in the sample carry out closet indexing, but rather to investigate the likely impact of closet indexing on investor outcomes among funds in our sample.

The continuous nature of the variables AS, TE and R^2 and their expected unidirectional relationship with closet indexing makes them individually well-suited to investigating the phenomenon. Beta, however, does not have unidirectional property; the incidence of closet indexing would be expected to be centred around values close to 1. This expectation is confirmed by the data (Chart 4).

In addition to examining the relationship of the variables taken individually with performance and costs, we also have reason to investigate how the variables together relate to these

outcomes. Interdependence of the variables (as analysed in section 5.1) and the non-linear expected dependence of closet indexing on beta would complicate a joint regression of the untransformed variables and its interpretation. A tractable way to address this problem is to combine different metrics in a single, binary variable. We do this with reference to methodology adopted by the Central Bank of Ireland in recent work. This approach allows us to test whether the metrics are jointly associated with directional effects on investor outcomes.

In common with the prevailing approach in the literature, including Cremers and Petajisto (2009) and Amihud and Goyenko (2013), we run pooled OLS regressions, including time fixed effects.[27] This approach, in contrast to a model including fund fixed effects, enables us to identify relationships among variables measured across entities, the intended focus of our analysis. As such, we identify differences in variables of interest (such as fund alpha) between potential closet indexers and the rest of the population of active funds, controlling for observed characteristics.

A model with fund fixed effects would have the advantage of controlling for unobserved characteristics that vary across funds. However, in our case, a plausible important unobserved characteristic whose effect we want to capture when comparing across funds is the effect of manager skill on performance. If, for example, higher-skill managers tend to adopt more active strategies, our estimates based on pooled OLS will capture this effect.[28]

In contrast with Cremers and Petajisto (2009) and Amihud and Goyenko (2013), we do not lag the explanatory variables, as our focus throughout this paper is on closet indexing metrics relate to current performance and costs.[29]

3.1 Relation between performance-based and portfolio-based metrics

We first test how the performance-based metrics TE, R^2 and beta relate to the portfolio-based metric AS. We define a combined indicator of potential closet indexing, denoted $potCI_{i,t}$, for fund i in year t, as follows.

$$potCI_{i,t} = \begin{cases} 1, & TE_{i,t} < 3\% \cap R^2_{i,t} > 95\% \cap 0.95 < \beta_{i,t} < 1.05 \\ 0 & otherwise \end{cases} \quad (1)$$

Recall from the discussion in section 2.1 that the incidence of closet indexers is expected to decrease in TE, increase in R^2 and to be greater in the region of beta values close to one.

The choice of threshold values for TE and R^2 is guided by ESMA (2016). In the case of the combined metric, the share of false positives among funds classified as potential closet indexers can be expected to decrease as the relevant thresholds are made stricter. Consequently, we also examine modified versions of the binary variable based on stricter criteria. Annex 1

[27] Amihud and Goyenko also include style fixed effects, encoded by a category variable in which each fund is identified with one of nine different management styles. Controlling for management styles would be a possible extension to our study for future work.

[28] A drawback is that we are then unable to control for other unobserved characteristics that may be correlated with observed independent variables. Nevertheless, varying the regression specifications to include fixed effects at fund level (in addition to the existing time dummies) yields comparable results in terms of the direction and significance of the closet indexing metrics in many cases. Differences are that in the regression of fund performance, the coefficients for AS and *potCI* lose significance in the presence of controls.

[29] Nonetheless, varying the regression specifications such that all independent variables are lagged by one year yields comparable results in terms of the directions and significance of almost all the closet indexing metrics, in all three of our main regressions. A difference was that lagged R^2 only predicts alpha at the 5% significance level in the presence of controls. An interpretation of the general results for the predictive performance specification is then that closet indexing metrics predict underperformance (as measured by net alpha). This predictive interpretation was a key focus in Amihud and Goyenko (2013). As a further robustness check, we then included lagged alpha as an additional independent variable. We found that this resulting in lagged R^2 no longer predicting alpha at the 10% significance level, though other closet indexing metrics remained significant predictors in the presence of controls. The reduction in significance of lagged R^2 when controlling for lagged alpha echoes Amihud and Goyenko's own findings in this regard.

illustrates average performance and costs for funds for which the binary variable equals one with those for which it equals zero under each of the specifications examined.

In choosing the threshold values for the core specification of the metric $potCI_{i,t}$ in equation (1), we therefore calibrate thresholds that are strict enough to allow for directional effects to be identified when we subsequently study how the combined metric relates to cost and performance, to complement our study of how the individual components of the metric relate to these outcomes. Reducing the threshold in R^2 to 90%, for example, prevents the combined metric from identifying a directional effect on returns, as illustrated by Charts 12 and 13 in Annex 1. A notable property of the metric is that the directional effects on investor outcomes are preserved as the metric is 'tightened' by making the thresholds stricter (Charts 12-14 in Annex 1).[30] This consistency property suggests that in the region of the joint distribution of the component metrics TE, R^2 and beta where false positives are sufficiently low to permit meaningful analysis, the inferred impact of potential closet indexing is qualitatively the same as that for the metrics studied individually. Furthermore, our key qualitative results do not appear sensitive to the choice of thresholds within this region of meaningful analysis, based on this informal check.

To test the extent to which the combined returns-based indicator explains AS in our sample, we carry out the following pooled OLS panel regression.[31]

$$AS_{i,t} = \beta_0 + \beta_1 potCI_{i,t} + \beta_2 X_{i,t} + \beta_3 W_{c,t} + \sum_{t=1}^{8} \delta_t T_t + u_{i,t} \quad (2)$$

where $AS_{i,t}$ is the AS for fund i at time t, $X_{i,t}$ is the set of fund-level characteristics and $W_{c,t}$ is the set of country-level characteristics.[32] In order to control for shocks whose impact is restricted to a given time period, time dummies T_t are included.

Controlling for these features, the β_1 coefficient is the average difference between the AS of potential closet indexers and funds that pursue active strategies. A positive sign would suggest that potential closet indexers have a higher AS on average, and conversely for a negative sign. Given the theoretical background in section 2, the predicted sign of β_1 in equation (2) is negative.

3.2 Closeting indexing and investor outcomes

We then move to the main focus of the paper, which is to examine the extent to which potential closet indexers are associated with higher or lower performance and costs. We carry out several regressions, specified as follows.

$$Y_{i,t} = \beta_0 + \beta_1 C_{i,t} + \beta_2 X_{i,t} + \beta_3 W_{i,t} + \sum_{t=1}^{8} \delta_t T_t + u_{i,t} \quad (3)$$

Where, according to the specification, $Y_{i,t}$ denotes (Jensen's) alpha or TER and $C_{i,t}$ denotes $AS_{i,t}$, $TE_{i,t}$, $R^2_{i,t}$ or $potCI_{i,t}$, where $TE_{i,t}$, $R^2_{i,t}$ are respectively TE and R^2 for fund i at time t and the other variables are defined as before. Since we are considering a yearly frequency, the

[30] Raising the R^2 threshold to 99% as in Charts 12-14 in Annex 1 illustrates this effect clearly, though at this point the number of funds identified by the metric falls to around 50, at which point inference from the data starts to become more difficult. The central choice of thresholds for the regressions in section 5 is thus in part a trade-off between discriminatory power and sample size.

[31] Recall from the discussion in section 2.1 that the incidence of closet indexers is expected to decrease in TE, increase in R^2 and to be greater in the region of beta values close to one.

[32] More precisely, the country-level characteristics vary according to fund domicile. All controls other than the time dummies are listed in the first column of each table of regression results in section 5.

employment of time dummies allows to control for shocks whose impact is restricted to a given time-period.

Based on the theoretical framework illustrated in Chart 1, AS should be negatively correlated with our composite performance-based indicator $potCI$, as the AS of a portfolio is the component where there is potential variability between fund and benchmark returns. On the other hand, the relationship between active management and performance seems to be less clear (Charts 8 and 9 illustrate the case of benchmark-relative performance). Predicting the sign of β_1 (for any choice of metric) in equation (3) when returns are the dependent variable is therefore less straightforward at this stage. If the theoretical framework of Sharpe (1991) approximately holds, we would expect the coefficient to be small in magnitude.

Finally, when TER is the dependent variable in equation (3), one would expect β_1 to be: (i) non-positive, assuming closet indexing is not related to pricing power; and (ii) small in magnitude. The latter hypothesis is based on the theoretical observations that setting significantly lower prices would lower the economic incentive to do closet indexing though some undercutting on price (facilitated by the fact closet indexers would bear lower economic costs than active strategies entail) may win market share.[33]

4. Data

4.1 Data description

We use yearly data over the period 2010 to 2018 for a sample of about 5,400 equity funds domiciled in the EU. Annual frequency is a natural choice given the long timeframe (9 years) and the typical reporting frequencies of many items on funds' balance sheets. In order to avoid any time discrepancy with yearly reported variables, we annualized data on excess (benchmark-relative) returns, TE, R^2 and beta.[34] Using annual frequency is beneficial for two main reasons. One, annual data are less volatile than lower-frequency data, allowing for relatively stable identification of potential closet indexers for the purpose of the present analysis. Two, annual frequency removes the risk that a fund could be erroneously identified as potential CI if TE, R^2 and beta are influenced by seasonal factors. However, a limitation is that data reflecting end-year positions may not be fully representative, if for example there are incentives for fund managers to meet end-of-year targets for performance or risk management by making temporary changes to their portfolios.

Selection of the sample is guided by the specification in ESMA (2016). The sample comprises UCITS equity funds, domiciled in the EU, that were not categorised as index-trackers and that had management fees of more than 0.65% of the total net asset value (NAV) of the fund.[35]

The measures of potential closet indexing included in our dataset are AS, TE, R^2 and beta, as defined in section 2.1.[36] While the latter three measures are available as reported by funds, AS

[33] A related theoretical constraint on optimal price-setting from the perspective of a closet indexer is that setting fees sufficiently low may reveal to the market that the manager faces lower economic costs than active strategies typically entail, thereby revealing that the fund is not genuinely active.

[34] These are calculated based on monthly returns

[35] Recall that the sample is composed of only funds that were declared to be active. We therefore extracted fund-level data from the Morningstar platform only for funds recorded as non-index-trackers.

[36] TE is therefore defined as the standard deviation of the absolute difference between fund and benchmark returns; R^2 and beta are based on a linear regression of fund returns on benchmark returns plus a constant. AS calculated against the primary prospectus benchmark return was not available. Instead, the AS used was calculated via the Morningstar terminal using Morningstar technical benchmarks.

is the result of a calculation which combines, at the fund level, a fund's portfolio and its benchmark index. Furthermore, the dataset includes many fund-specific characteristics such as fund size and age, returns (net of costs and gross, benchmark-adjusted and unadjusted), alpha and TER. Finally, the dataset includes macro-level data that capture the macroeconomic and market environment over time such as inflation and market volatility.

AS is calculated against technical benchmarks assigned by the data provider, Morningstar. As such, the measure of AS we use in the reported regressions does not necessarily reflect how far a fund's holdings differ from those in the benchmark it reports in its prospectus. An advantage of using AS based on technical benchmarks is that it has higher coverage; the sample size would be around one quarter lower if we were to use AS based on the prospectus benchmark. However, a disadvantage is that unlike the prospectus benchmark, the technical benchmark does not form part of the information disclosed to investors.[37]

The micro-level data originate from three commercial data terminals. Data on costs and net returns are from Refinitiv Lipper, as reported by funds. All other figures on fund characteristics and performance are from Morningstar Direct, as reported by funds. Finally, the set of macroeconomic data (inflation, VSTOXX, etc) is retrieved from Refinitiv Eikon.

One issue encountered in constructing the dataset was missing observations. All those funds with missing values of AS, TE and R^2 were excluded from the sample. Another issue was that some outliers were detected that took implausible values. For example, AS for some funds vastly exceeded 100%, even though none of the funds in the sample were heavily leveraged. Funds whose data included such outliers were dropped from the sample. Following deletion of missing or suspect observations, the final sample amounted to 3,206 UCITS funds. The final sample has a total size of EUR 1.41tn.[38]

Most variables are expressed in percentage form. In order to guarantee some degree of stationarity, trending variables such as VSTOXX have been transformed in first differences.

A possible limitation in our dataset is that is does not allow us to identify the extent to which fund managers may be using derivatives to take active positions (in which case, activeness would be underestimated by our metrics) or to hedge them (in which case, activeness would be overestimated).[39]

4.2 Market overview

In this subsection we present the descriptive information regarding the different metrics we use to study closet indexing. We then split the sample into two groups according to our composite indicator of potential closet indexing and report the evolution of fund-specific characteristics (fund size, age, alpha and TER) over time for each identified group.

[37] As a robustness check, we ran our regressions using AS calculated against prospectus benchmarks. The main difference was in the case of the performance (alpha net of costs) regression, where AS was significant at the 5% level in the absence of controls but lost significance in the presence of controls. However, AS was significantly associated (at the 1% level) with performance measured by gross returns

[38] For comparison, according to data published by EFAMA (2019), total AuM in equity UCITS was around EUR 3.5tn at the end of 2018. EFAMA does not split the data by whether a fund's mandate is active or passive. However, as reported in the ESMA Report on Trends, Risks and Vulnerabilities No. 2, 2019, p. 60, around a quarter of AuM in equity UCITS is estimated to be in funds with passive mandates.

[39] Another possible limitation we considered was funds may follow 'smart beta' strategies. Such strategies may involve tracking a bespoke benchmark that in effect carries out a pre-specified strategy, such as maintaining certain factor weights in the portfolio. Such cases could in theory complicate our analysis. However, while our dataset includes a variable denoted whether a fund uses smart beta strategies, none of the funds in our sample were reported as doing so.

Table 1
Summary of key variables
Descriptive statistics for CI metrics and fund characteristics

	Obs.	Mean	Min	Max	St. Dev	P10	P25	P50	P75	P90
CI metrics										
Beta	25,426	1.0	-1.2	3	0.2	0.8	0.9	1	1.1	1.2
R^2	25,426	82.9	0	100	16.7	60.8	77.1	88.4	94.5	97.5
AS	25,889	76.9	9.1	100	18	50	67.6	80.7	90.9	96.7
TE	25,426	5.1	0.1	35.2	2.9	2.2	3.2	4.5	6.3	8.6
Fund characteristics										
Fund size	28,683	443	0	13,100	914	12.4	41.9	143	423	1,120
Net flows	26,093	1.3	-7,790	5,230	216	-90.1	-22.4	-1.5	14.5	91.6
Alpha	24,423	9.7	-65.9	138.3	15.2	-11.2	-0.6	10.8	20.2	27.6
TER	24,002	1.7	0	10	0.6	1.0	1.2	1.7	1.9	2.3
Age	26,168	13.1	0	85.0	9.8	2	6	12	18	25

Note: R^2, AS, TE, alpha and TER in percentage points. "Alpha" = Jensen's alpha for a fund at year-end based on a 36-month trailing calculation. Fund size and net flows in EUR mn. Age in years. "Obs." = Total observations. "P10"=Value of variable at top of first decile; subsequent columns analogously defined.

First, we check whether the components of our potential closet indexing indicator – TE, R^2 and beta – relate to AS as predicted. Recall that AS is a measure of active management that can be interpreted as the fraction of the portfolio that is different from the index.[40] As closet indexers aim to track the benchmark index, their AS should be low. Return-based measures such as TE, R^2 and beta should largely be governed by the active component of a fund portfolio.

Chart 2 shows positive correlation between AS and TE, as expected based on the theoretical framework and empirical results summarised in section 2. In other words, funds whose returns change in line with changes in benchmark returns tend to be those with a small active component in their portfolio. Chart 3 shows negative correlation between AS and R^2. Intuitively, funds whose past performance is largely explained by benchmark performance have a lower active share. Other things equal, the higher the R^2, the higher the chance that a fund is a closet indexer.

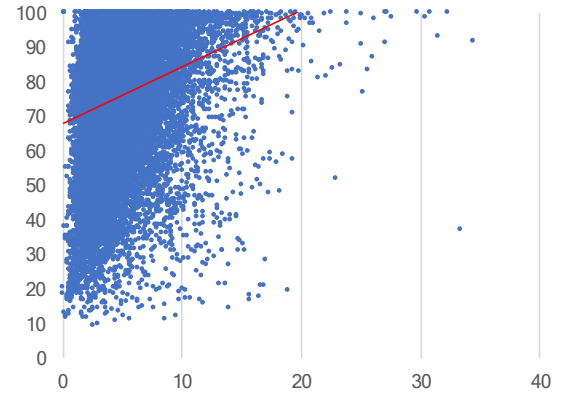
Unlike TE and R^2, beta has a non-linear relationship with AS (Chart 4). This arises from the fact that beta measures benchmark exposure. A value of one means that it replicates the fund but significant deviations in either direction indicate that a given fund is not close to replicating its benchmark.

[40] See Cremers and Petajisto (2009)

Chart 2

Tracking Error vs. Active Share

Positive relationship between AS and TE

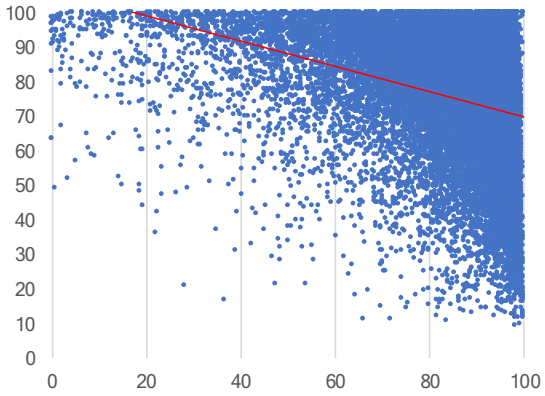


Note: Scatter plot between TE (x-axis) and AS (y-axis). Trendline in red.
Sources: Morningstar Direct, ESMA

Chart 3

R^2 vs. Active Share

More active funds have lower R^2

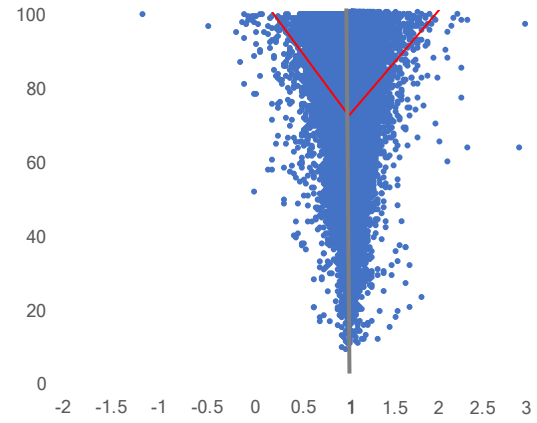


Note: Scatter plot between R2 (x-xis) and AS (y-axis). Trend line in red.
Sources: Morningstar Direct, ESMA

Chart 4

Beta vs. Active Share

Discontinuous trend around 1

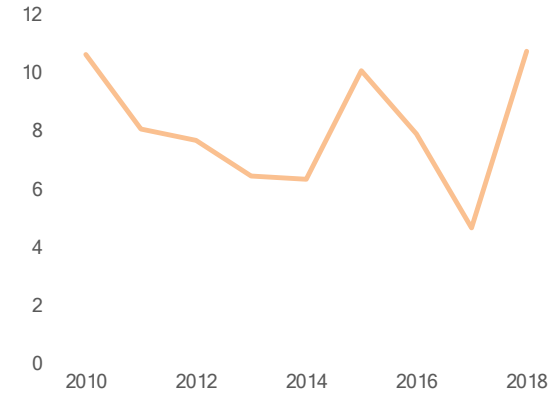


Note: Scatterplot of beta (x-axis) and AS (y-axis) for funds sampled.
Sources: Morningstar Direct, ESMA

Chart 5

Share of funds satisfying *potCI* criteria

Some more variation in recent years



Note: Share of equity funds from our sample that satisfy the PotCI indicator variable in a given year.
Sources: Morningstar Direct, ESMA.

Charts 2 to 4 illustrate that the data are consistent with the hypothesis – motivated by the theoretical and empirical literature – that funds engaging in passive strategies have low TE, high R^2, and beta that diverges from one. In particular, the data suggest that the CI indicator thresholds we use in our central case are sufficiently congruent with the direction of the plots.

The *potCI* indicator – which as explained in section 3.1 takes into account only returns-based metrics rather than AS – shows some variation over time (Chart 5).[41] We find that a large

[41] The reason for using *potCI* in the present context is to enable tractable study of the joint relation of several closet indexing metrics with outcome variables of interest. It is not intended to definitively identify potential closet indexers at a fund level, and ultimately identification of closet indexers will rest on supervisory work such as granular analysis of fund prospectuses and KIIDs. See the discussion of equation (1) in section 3 for further details.

majority of funds rarely change their potCI value over time, with over 70% not changing at all over the nine annual observations. However, 6.5% of funds change status three or more times over the time series. We posit that these funds are 'marginal', in the sense that they have values of one or more of the components of the potCI metric that are close to relevant thresholds. As such, their potCI status is quite likely to change based on market conditions in a given year.

An alternative hypothesis would be that some fund managers may decide to take more active positions in some years than in others. One possibility is that during periods of heighted uncertainty or market volatility (such as before the sovereign debt crisis in 2011-12 and before the June 2016 UK referendum), the incentive for closet indexing may increase, as there is greater scope for poor relative performance. For 2018, the most recent year in the sample, potential closet indexing as measured by the combined metric covers is at a peak of almost 11% of the sample. Uncertainty around international trade, for example, may have induced funds to pursue "safer" strategies and hug the index. Another possible factor affecting the economic incentive for closet indexing is monetary policy. One possibility is that expansive policy may increase market returns as a whole, while at the same time increasing correlation among equity returns, thus lowering the relative returns from stock-picking compared to exposure to the whole market.[42]

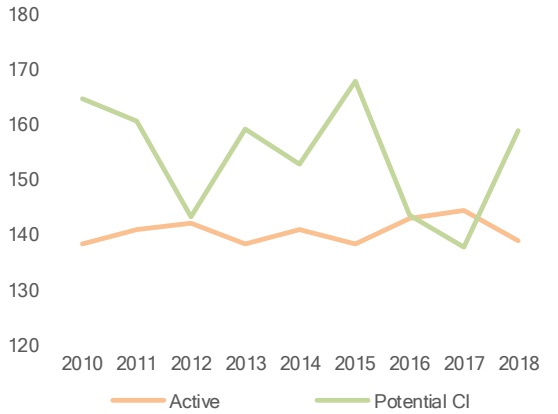
Another possibility is that investing in large cap companies offers fewer opportunities for profitable active strategies, if larger companies are more widely researched than smaller ones. We find that funds in our sample investing in large cap companies tend to have a slightly higher incidence of *potCI* than funds investing in smaller companies.

For the distribution of the individual metrics TE, R^2, beta and AS over the whole sample of funds by year, see charts 16-19 respectively in Annex 1.

Chart 6

Fund size by *potCI* value, 2010-2018

Potential CIs tend to be larger

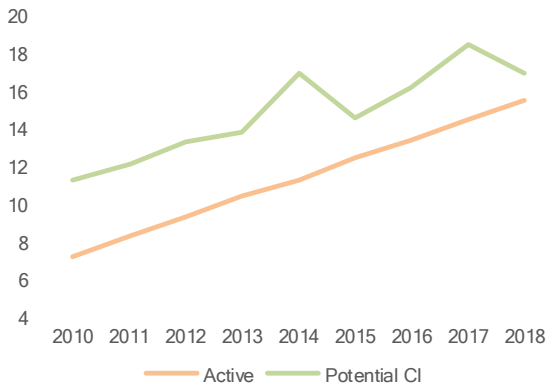


Note: Evolution of median fund size in EUR mn over time of the two identified groups. "Potential CI"=funds satisfying the Pot CI indicator variable in a given year.
Sources: Morningstar Direct, ESMA

Chart 7

Fund age by *potCI* value, 2010-2018

Younger funds tend to be more active



Note: Average age in years (y-axis) over time (x-axis) of the two identified groups. "Potential CI"=Funds satisfying the PotCI indicator variable in a given year, "Active"=other funds in the sample.
Sources: Morningstar, ESMA

[42] For examples of this view as expressed by industry participants, see Levy (2017).

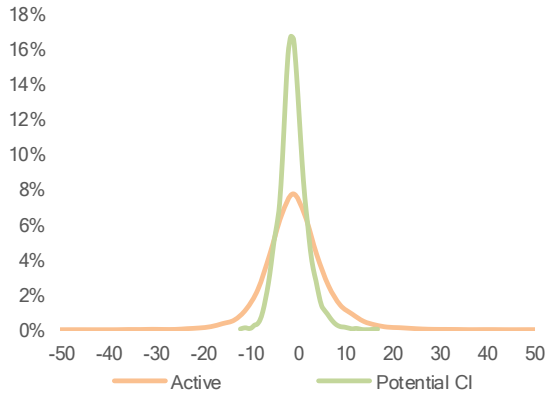
Turning to the evolution of fund size over time (Chart 6), the median fund that meets our combined metric for potential closet indexing is larger than that for the result of the population in all but one of the years sampled. This result echoes Cremers and Petajisto (2009), who find a negative correlation between AS and fund size. Chart 7 shows that on average, younger funds tend to be more active as measure by our combined metric, in line with findings in Pastor et al (2014), who offer the explanation that younger funds are more likely to pursue the latest strategies, while older funds tend to innovate less.

Excess returns, defined as gross fund returns minus gross benchmark returns, are a variable of interest as managers' performance is often judged relative to that of the market in which they invest. In keeping with the literature, however, we use alpha as our primary measure of performance as it also adjusts for the level of risk borne by a manager, as proxied by benchmark exposure. Both excess returns and alpha for funds that meets our combined metric for potential closet indexing show far less variability than the remainder of the population (Charts 8 and 9). This pattern is unsurprising, as TE, R^2 and beta all represent ways in which fund returns can be said to differ from benchmark returns. Intuitively, 'hugging the benchmark' results in lower variation in excess returns. The means of the distributions in Charts 8 and 9 appear slightly different, with expected returns slightly lower for potential closet indexers in both cases. This informal result is confirmed by regression analysis of alpha via equation (3), as reported in Table 3. Regression analysis of excess returns gives similar confirmation.

Chart 8

Excess return distributions by *potCI* value

Concentrated performance for potential CIs

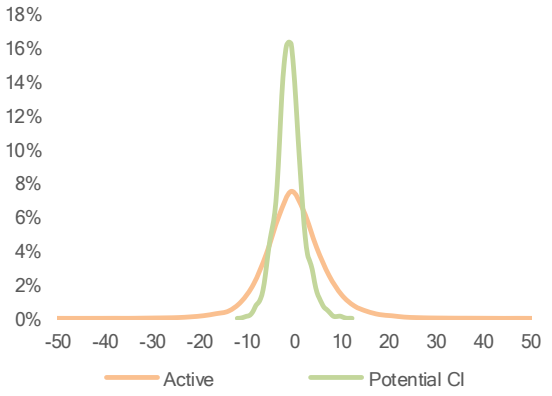


Note: Kernel density in % (y-axis) of excess returns (x-axis) of the two identified groups. "Potential CI"=Funds satisfying the PotCI indicator variable in a given year, "Active"=other funds in the sample.
Sources: Morningstar Direct, ESMA

Chart 9

Jensen's alpha distributions by *potCI* value

Concentrated performance for potential CIs

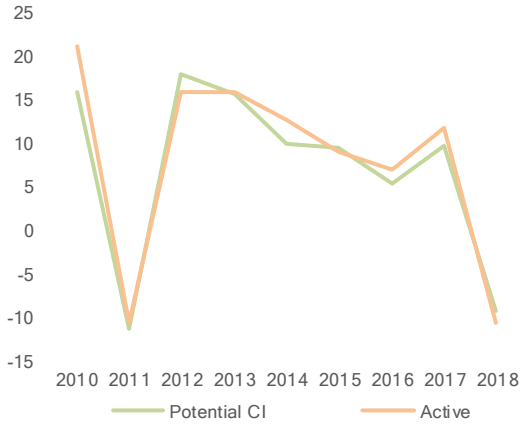


Note: Kernel density in % (y-axis) of Jensen's alpha (x-axis) of the two identified groups. "Potential CI"=Funds satisfying the PotCI indicator variable in a given year, "Active"=other funds in the sample.
Sources: Morningstar Direct, ESMA

Chart 10

Jensen's alpha by *potCI* value, 2010-2018

Slightly lower performance for potential CIs

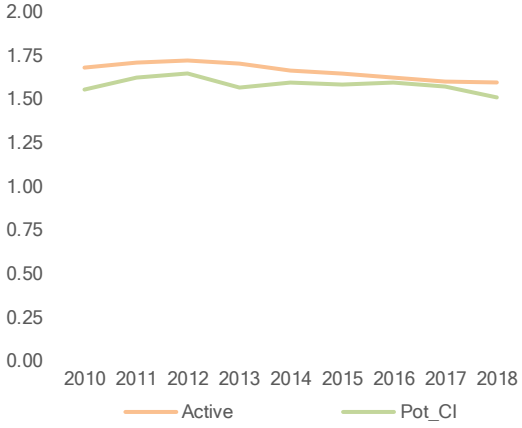


Note: Evolution of average Jensen's alpha, % (y-axis) over time (x-axis) of the two identified groups
Sources: Morningstar Direct, ESMA

Chart 11

TER by *potCI* value, 2010-2018

Potential CI barely any cheaper



Note: Evolution of average Total Expense Ratio of the two identified groups, percentage points (y-axis).
Sources: Refinitiv Lipper, Morningstar DIrect, ESMA.

On average, funds in our sample not classified as potential closet indexers performed moderately better in gross terms than potential closet indexers, not adjusting for benchmarks or other factors (Chart 10), also confirmed by regression analysis. As noted in the discussion of Chart 5, several factors are likely to have driven (unadjusted) gross performance in the years in the sample. One reason for the large fall in gross performance in 2011 is market perceptions of contagion risk from the European sovereign debt crisis. The 2018 drop in equity returns may relate to elevated economic and political uncertainty.[43]

As expected, average TER stays almost constant across time (Chart 11), though with a slight downward trend, and with potential closet indexers showing slightly lower TER.[44] As noted in section 2.2, closet indexers may have some incentive to lower their fees to win market share, provided doing so does not indicate to the market that they are not in fact pursuing active strategies. However, to maximise total fund manager revenue, they need to balance this against the direct impact of foregone revenue from cutting fees.

5. Empirical results

In this section we set out the regression results of the models specified in section 3. In section 5.1, we examine how the performance-based metrics TE, R^2 and beta relate to the portfolio-based metric of AS among the equity funds in our sample. In section 5.2, we examine how the metrics of potential closet indexing relate to performance and costs. Both sets of regressions include controls for macroeconomic environment and fund characteristics. For the reasons set

[43] For further explanations please refer to the ESMA Report on Trends, Risks and Vulnerabilities No. 1, 2019.

[44] The slight downward trend in TER across the population of active funds as a whole may be due to increasing competitive pressure from passive funds, as documented for example in Cremers et al (2016).

out in section 3, we run unlagged pooled OLS regressions including time fixed effects. Varying our specifications respectively to include fund fixed effects, lagged independent variables and replacing alpha with other performance metrics yields qualitatively similar results in many cases.[45]

5.1 Determinants of Active Share

To better understand the determinants of AS, we estimate panel regressions where the dependent variable is the yearly fund-level AS and the variable of interest is our combined returns-based measure of potential closet indexing, as described in section 4.2. We control for fund characteristics and macroeconomic factors. We introduce year dummies to take into account fixed effects within years. Standard errors are clustered by fund.

Table 2
Regression results
Determinants of Active Share

	(1) Active Share	(2) Active Share	(3) Active Share	(4) Active Share	(5) Active Share	(6) Active Share	(7) Active Share
potCI	-17.915*** (0.000)	-15.996*** (0.000)	-15.931*** (0.000)	-15.901*** (0.000)	-13.970*** (0.000)	-13.873*** (0.000)	-12.794*** (0.000)
Age		-3.230*** (0.000)	-3.447*** (0.000)	-3.197*** (0.000)	-3.946*** (0.000)	-3.942*** (0.000)	-4.116*** (0.000)
Fund value		-9.189*** (0.000)	-8.938*** (0.000)	-9.267*** (0.000)	-9.838*** (0.000)	-9.799*** (0.000)	-8.918*** (0.003)
(Fund value)^2		0.213*** (0.000)	0.209*** (0.000)	0.215*** (0.000)	0.230*** (0.001)	0.229*** (0.001)	0.209*** (0.005)
Market volatility		0.047 (0.104)	0.108*** (0.002)	0.042 (0.144)	0.069 (0.1)	0.065 (0.121)	0.098* (0.068)
Inflation		-0.469 (0.105)	-0.159 (0.642)	-0.526* (0.07)	-0.462 (0.221)	-0.499 (0.188)	-0.269 (0.571)
Net returns			0.125*** (0.000)				0.102*** (0.000)
Net flows				0.000** (0.023)		0.000 (0.297)	0.000 (0.454)
Net expenses					1.808*** (0.001)	1.798*** (0.001)	2.097*** (0.000)
N	25889	21840	17874	21709	11013	10986	9119
N cluster	2916	2903	2541	2897	2490	2485	2098
R^2	0.075	0.103	0.111	0.102	0.094	0.093	0.093

Note: Annual observations from 2010 to 2018. "Market volatility"= the annualized daily VSTOXX. Inflation is yearly HCPI. “Net returns” = fund returns net of costs, not benchmark-adjusted. “N cluster” = number of observations when clustered by fund ID. All specifications include time dummies and a common intercept. Standard errors clustered by fund ID. P-values in parentheses; ***p<0.01, **p<0.05, *p<0.1

[45] See the discussion at the start of section 3, including footnotes 27-29, for further details.

As shown in Table 1, the returns-based indicators of potential closet indexing together have a significant negative effect on AS across all specifications. Funds that meet the conditions specified in equation (1) have an AS around 15 percentage points higher than those not meeting the conditions. While the explanatory power of the regression as measured by the goodness of fit (reported in the final row of Table 1) is low, the persistence of the results across different specifications suggests that the CI indicator is a robustly significant predictor of AS in our sample.

The results in Table 1 also suggest that certain fund characteristics play a role in determining AS. In particular, the age of a fund appears to influence AS, with older funds on average less active than younger funds. Higher net performance is associated with slightly lower AS, while more active funds seem to be on average significantly more expensive than less active funds, once we control for net performance. We include a quadratic term in fund size among the controls given the finding by Cremers and Petajisto (2009) that the relationship between AS and fund size is non-linear (specifically, that AS is roughly constant among smaller funds but after a certain threshold, decreases in fund size). Our results confirm a similar non-linear relationship when all controls are included (column 7 of Table 1).

In summary, our combined performance-based metric for potential closet indexers is a strong predictor of AS. Various fund characteristics such as age, and fund size are associated with AS in a manner in line with the existing literature. We find that higher net (unadjusted) performance is associated with slightly higher AS.

5.2 Determinants of performance and costs

In this section we analyse the effect of potential closet indexing variables, both individually and via a combined measure, on fund performance and costs. As in section 5.1, we control for fund characteristics and macroeconomic factors, and cluster standard errors by funds.

Table 3
Regression results
Impact of potential closet indexing on alpha

	(1) Alpha	(2) Alpha	(3) Alpha	(4) Alpha	(5) Alpha	(6) Alpha	(7) Alpha	(8) Alpha
Tracking Error	0.222***	0.198***						
	(0.000)	(0.000)						
R^2 [fund variable]			-0.076***	-0.072***				
			(0.000)	(0.000)				
potCI					-1.106***	-0.880***		
					(0.000)	(0.000)		
Active Share							0.018***	0.013***
							(0.000)	(0.000)
Age		-0.246***		-0.192***		-0.262***		-0.234***
		(0.000)		(0.001)		(0.000)		(0.000)
Fund value		-0.27		0.15		-0.51		-0.48
		(0.62)		(0.787)		(0.338)		(0.366)
(Fund value)^2		0.01		0.00		0.01		0.01
		(0.553)		(0.898)		(0.32)		(0.344)

Net flows		0.000*** (0.000)		0.000*** (0.000)		0.000*** (0.000)		0.000*** (0.000)
Market volatility		-0.028*** (0.005)		-0.055*** (0.000)		-0.032*** (0.002)		-0.031*** (0.003)
Inflation		0.256*** (0.000)		0.210*** (0.001)		0.275*** (0.000)		0.299*** (0.000)
N	22908	21254	22908	21254	22908	21254	22908	21254
N cluster	2863	2847	2863	2847	2863	2847	2863	2847
R^2	0.093	0.114	0.119	0.137	0.087	0.109	0.088	0.109

Note: Annual observations from 2010 to 2018. "R^2 [fund variable]" = coefficient of determination obtained from CAPM regression of fund returns on benchmark returns which is then used as an independent variable in regressions (3) and (4). The R^2 value in the final row is the coefficient of determination from the regressions reported in the table. "Alpha" = Jensen's alpha for a fund at year-end based on a 36-month trailing calculation. "Market volatility" = annualised daily VSTOXX. Inflation is yearly HCPI. "Net returns" = fund returns net of costs, not benchmark-adjusted "N cluster" = number of observations when clustered by fund ID. All specifications include time dummies and a common intercept. Standard errors clustered by fund ID. P-values in parentheses; ***p<0.01, **p<0.05, *p<0.1.

Table 3 reports the results of panel regressions of alpha on the different proxies for potential closet indexing. The coefficients of the different potential closet indexers measures are strongly significant in all specifications. Furthermore, the sign of the coefficients consistently show that potential closet indexers tend to have significantly lower performance. This result is line with results from the literature, discussed in section 2.2, that more active funds tend to have higher excess returns on average.[46] With reference to Chart 1 in section 2.1, the fact that this relationship is consistently observed across portfolio-based and returns-based measures suggests that different approaches to active management – including factor bets, diversified stock picks and correlated stock picks – are positively associated with higher performance among the equity funds in our sample.[47]

Alpha is a popular measure of fund performance in many studies as it adjusts returns for the market risk of a fund, as measured by benchmark exposure. A simpler approach is to look at unadjusted returns. We find that the results in Table 3 are qualitatively similar when unadjusted gross returns are used as the dependent variable instead of alpha. The same is true for excess returns measured simply as the difference between net returns and benchmark returns.[48]

Table 4 reports the results of panel regressions of TER over different proxies for potential closet indexing.[49] The results suggest that potential closet indexers are slightly cheaper than the wider population of truly active funds. Significant effects can be found via return-based metrics for potential closet indexing both in the single and combined regressions. Specifically, TE, R^2 and *potCI* are valid predictors at the 1% significance level. In terms of magnitude, potential closet indexers are 0.06 pp (i.e. 6 bp) cheaper than truly active funds. Similarly, funds with higher tracking error (truly active) have a higher TER on average. The effect of a 1pp decrease in R^2 is associated with a decrease in TER of less than a tenth of a basis point.

As set out in Box 3, TER captures many costs. Some of the (very small) difference in average TER by PotCI value is attributable to management costs for example, which are a significant

[46] As discussed in section 2.2, if the aggregate portfolio of funds that adopt active strategies differs from the market, there is scope for performance to be associated with greater activeness.

[47] Recall from section 2 that higher values of TE, *potCI* and AS indicate a greater degree of active management, whereas lower values of R^2 indicate a greater degree of active management.

[48] An exception is that TE is significantly associated with gross returns only at the 10% level without the control variables listed in Table 3, and the 5% level when the controls are included.

49 We also examined the effect of augmenting the TER by including a proxy for incurred load and unload fees, based on fund flows. The methodology is the same as used in ESMA (2020). The regressions of costs give qualitatively similar results as in the case of the TER metric.

component of TER (Chart 21). At the same time, TER does not capture all possible costs. Notably, front-end and back-end loads, which can be significant costs to investors, are omitted. However, in qualitative terms the pattern of average costs by PotCI value is again similar (Chart 21). Likewise, regression results are similar if the cost measure is extended to include not only TER but also a proxy for front-end and back-end loads.[50]

Table 4
Regression results
Impact of potential closet indexing on TER

	(1) TER	(2) TER	(3) TER	(4) TER	(5) TER	(6) TER	(7) TER	(8) TER
Tracking Error	0.019***	0.017***						
	(0.000)	(0.000)						
R^2 (fund variable)			-0.001***	-0.001**				
			(0.007)	(0.021)				
potCI					-0.055***	-0.075***		
					(0.004)	(0.000)		
Active Share							0.000	0.001*
							(0.485)	(0.079)
Age		0.147***		0.146***		0.144***		0.145***
		(0.000)		(0.000)		(0.000)		(0.000)
Fund value		-0.337***		-0.355***		-0.355***		-0.354***
		(0.001)		(0.000)		(0.000)		(0.000)
(Fund value)^2		0.007***		0.007***		0.007***		0.007***
		(0.005)		(0.003)		(0.003)		(0.003)
Net flows		0.000		0.000		0.000		0.000
		(0.413)		(0.441)		(0.43)		(0.449)
Market volatility		-0.002**		-0.003**		-0.003***		-0.003**
		(0.032)		(0.01)		(0.008)		(0.014)
Inflation		-0.014		-0.012		-0.013		-0.011
		(0.148)		(0.234)		(0.193)		(0.276)
N	20935	19884	20935	19884	21695	20327	21695	20327
N cluster	2698	2686	2698	2686	2749	2735	2749	2735
R^2	0.015	0.087	0.009	0.082	0.008	0.081	0.007	0.081

Note: Annual observations from 2010 to 2018. "R^2 [fund variable]" = coefficient of determination obtained from CAPM regression of fund returns on benchmark returns which is then used as an independent variable in regressions (3) and (4). The R^2 value in the final row is the coefficient of determination from the regressions reported in the table. TER in percentage points. Inflation is yearly HCPI. "Market volatility" is the annualized daily VSTOXX. All specifications include time dummies. P-values in parentheses; ***p<0.01, **p<0.05, *p<0.1..

0.015 0.

[50] The methodology for estimate front-end and back-end loads is based on measures of fund flows, and is that used in ESMA (2020), which provides details.

In summary, the results in this section suggest that potential closet indexers on average show lower alpha than other active funds. This means that although there is evidence that potential closet indexers are slightly cheaper than truly active funds, the negative marginal effect of being a potential CI, or less active fund, seems to be prevailing for returns rather than for costs. In other words, not only do less active funds achieve worse performance, but they also tend to be almost as expensive as truly active funds.

6. Conclusion

Closet indexing can be defined as a practice whereby asset managers claim to manage their funds in an active manner while in fact tracking or remaining close to a benchmark index. As ESMA has previously highlighted, closet indexing is a major investor protection concern in its own right, as it involves misrepresenting information to investors.

We investigate how potential closet indexing – as measured by a range of different metrics – relates to performance and costs of EU equity funds. We find evidence that the potential closet indexing metrics we study are associated with lower alpha. This result is in line with several recent studies of US equity funds. Similar results hold for simpler performance measures such as unadjusted returns. Turning to costs, we find that potential closet indexers are associated with a slightly lower TER than active funds generally. Although closet indexing funds enjoy much lower economic costs than other active funds, they only pass on a small proportion of these savings to consumers on average, rather than competing strongly on price to win market share.

A possible topic for future work, building on the present study, would be to broaden the set of closet indexing metrics used still further, for instance by including return-based metrics calculated via a 4-factor model (thereby controlling for the known factors of size, value and momentum) or the Style-Shifting Activity metric.

In summary, our results suggest that outcomes for investors in closet indexing funds are on average worse than outcomes for investors in genuinely active funds. Investors face lower expected returns from closet indexers than from what they are promised, namely an actively managed fund portfolio. In other words, in addition to representing a form of misconduct in its own right, closet indexing investors makes worse off ex-ante. Even though potential closet indexers are marginally cheaper than genuinely active funds, this difference is outweighed by reduced performance: potential closet indexers perform worse even when fees are taken into account. More generally, our results provide strong confirmation of the concerns of supervisors and investor advocacy groups that investors in closet indexing funds face an unjustifiably high level of costs, far in excess of those for explicitly passive funds.

References

Amihud, Y. and R. Goyenko, 2013, *Mutual Fund's R^2 as a Predictor of Performance*, Review of Financial Studies, 26 (3), pp. 667-694

Better Finance, 2019a*, Better Finance Welcomes ESMA's Efforts at Rooting Out Closet Funds*, website article, 8 April

Better Finance, 2019b*, Study on the Correlation between Cost and Performances in EU Equity Retail Funds*, 20 June

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

Central Bank of Ireland, 2019, *Thematic Review of Closet Indexing*, Press Release, 19 June

Cremers, K. and Q. Curtis, 2016, *Do Mutual Fund Investors Get What They Pay For? The Legal Consequences of Closet Index Funds*, Virginia Law & Business Review 11(1), pp. 31-93

Cremers, K., M. A. Ferreira, P. Matos and L. Starks, 2016, *Indexing and active fund management: International evidence*, Journal of Financial Economics, 120, pp. 539-560

Cremers, K. and A. Pareek, 2016. *Patient capital outperformance: The investment skill of high active share managers who trade infrequently*, Journal of Financial Economics, 122(2), pp. 288-306

Cremers, K. and A. Petajisto, 2009, *How Active Is Your Fund Manager? A New Measure That Predicts Performance*, Review of Financial Studies, 22(9), pp. 3329-3365

Demartini, A. and N. Mosson, 2018, *Closet Index Funds: A Contribution to the Debate in Europe*, Autorité des Marchés Financiers

EFAMA, 2019, *Quarterly Statistical Release No. 76, Trends in the European Investment Fund Industry in the Fourth Quarter of 2018 & Results for the Full Year of 2018*, March

ESMA, 2016, *Public Statement on Potential Closet Indexing*, 2016/165, 2 February

ESMA, 2019, *Questions and Answers: Application of the UCITS Directive*, ESMA34-43-392, 4 June

ESMA, 2020, *Performance and Costs of Long-Term Investment Products in the EU*, ESMA Annual Statistical Report, ESMA 50-165-1098, 6 April

Fama, E. F. and K. R. French, 1993, *Common risk factors in the returns on stocks and bonds*, Journal of Financial Economics, vol. 33, pp. 3-56

Fama, E. F. and K. R. French, 2010, *Luck versus Skill in the Cross-Section of Mutual Fund Returns*, Journal of Finance, 65(5), October, pp. 1915-1947

Fong, W. M., 2016, *On the Impact of Closet Indexing in Active Fund Management*, Journal of Performance Measurement, pp. 27-32

Frazzini A., J. Friedman and L. Pomorski, 2016, *Deactivating Active Share*, Financial Analyst Journal, 72(2), 2016

Grinold, R. C. and R. N. Kahn, 1999, *Active Portfolio Management: A Quantitative Approach for Producing Superior Returns and Selecting Superior Returns and Controlling Risk*. McGraw-Hill Library of Investment and Finance

Jegadeesh, N. and S. Titman, 1993, *Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency*, Journal of Finance, 48(1), pp. 65-91

Kjørven, M., 2019, *Closet Index Funds and Retail Investor Protection – A Scandinavian Perspective*, Tilburg Law Review, 24 (1), pp. 125-138

Levy, D., 2017, *Will 2017 Favour More Active Fund Managers?* Morningstar Fund Research and Insights, 10 February

Morningstar Research, 2016, *Active Share in European Equity Funds*, April

Morningstar Research, 2019, *Morningstar's Active/Passive Barometer*, August

Pastor, L., R. F. Stambaugh and L. Taylor, 2015, *Scale and Skill in Active Management, Journal of Financial Economics*, 116(1), pp. 23-45

Petajisto, A., 2013, *Active Share and Mutual Fund Performance*, Financial Analysts Journal, 69(4), pp 73-93

Sharpe, W. F., 1991, *The Arithmetic of Active Management*, Financial Analysts Journal, 47(1), January/February, pp. 7-9

Annex 1: Additional descriptive statistics

Chart 12

Gross returns by *potCI* for different threshold criteria

Lower gross returns for stricter criteria

Note: Average annualised gross returns for funds within the sample for which a given set of criteria are satisfied (=1) or not (=0), %.
Sources: Morningstar Direct, ESMA.

Chart 13

Net returns by *potCI* for different threshold criteria

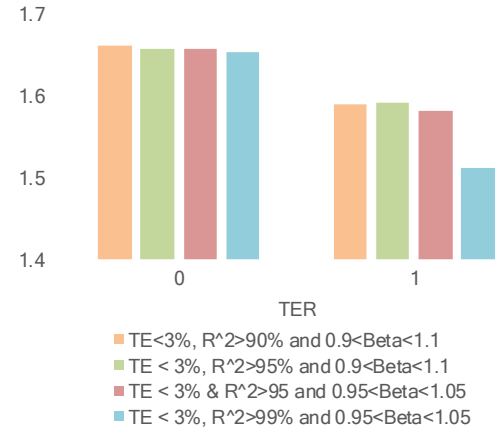
Lower net returns for stricter criteria

Note: Average annualised net returns for funds within the sample for which a given set of criteria are satisfied (=1) or not (=0), %.
Sources: Morningstar Direct, ESMA.

Chart 14

TER by *potCI* for different threshold criteria

Lower TER for stricter criteria



Note: Average TER for funds within the sample for which a given set of criteria are satisfied (=1) or not (=0), percentage points.
Sources: Morningstar Direct, ESMA.

Chart 15

Jensen's alpha by *potCI* for different threshold criteria

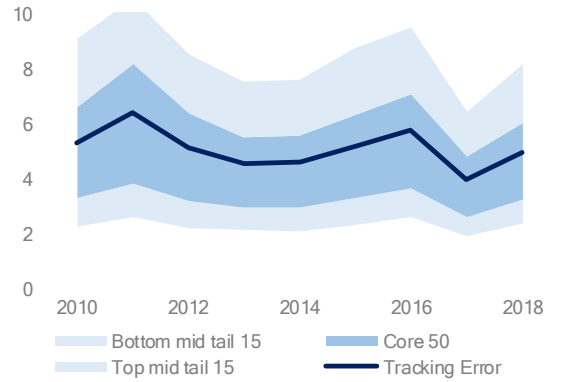
Negative alpha for potential CIs across criteria

Note: Average Jensen's alpha for funds within the sample for which a given set of criteria are satisfied (=1) or not (=0), %.
Sources: Morningstar Direct, ESMA.

Chart 16

Distribution of TE over time

TE lowest in 2017

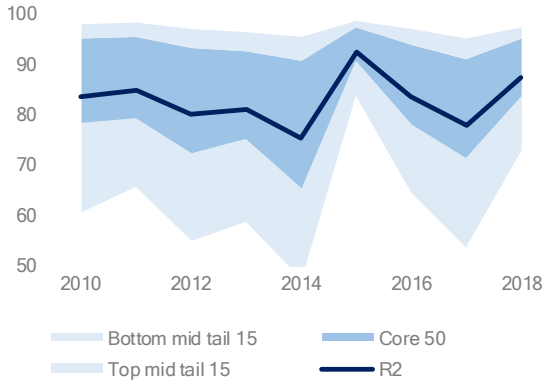


Note: "Tracking Error" = average value of fund TE in the sample. "Top mid-tail 15" =distribution between the 75th and 90th percentile. "Bottom mid-tail 15" =distribution between the 10th and 25th percentile.
Sources: Morningstar Direct, ESMA.

Chart 17

Distribution of R^2 over time

R^2 higher in 2015 and 2018

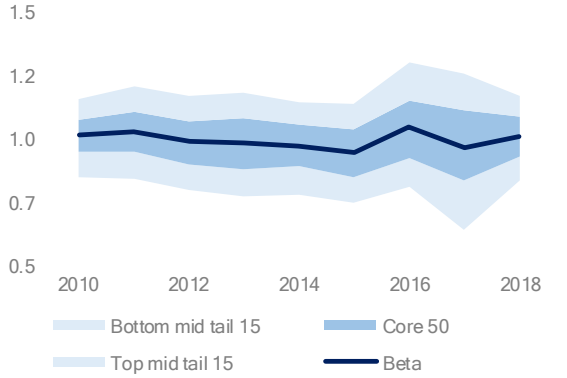


Note: "R2" represents the average value of fund R2 in in the sample. "Top mid-tail 15" =distribution between the 75th and 90th percentile. "Bottom mid-tail 15" =distribution between the 10th and 25th percentile.
Sources: Morningstar Direct, ESMA.

Chart 18

Distribution of beta over time

Beta highest in 2016

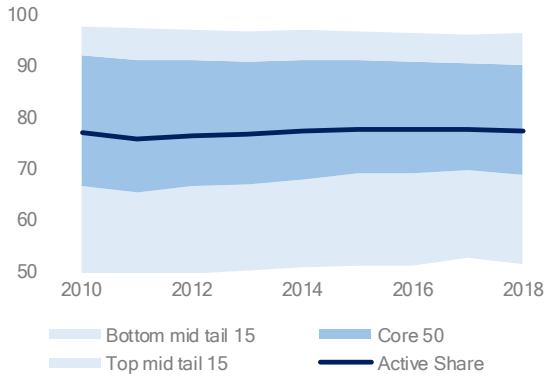


Note: "Beta" = average value of fund beta in the sample. "Top mid-tail 15" =distribution between the 75th and 90th percentile. "Bottom mid-tail 15" =distribution between the 10th and 25th percentile.
Sources: Morningstar Direct, ESMA.

Chart 19

Distribution of AS over time

High degree of stability in mean AS

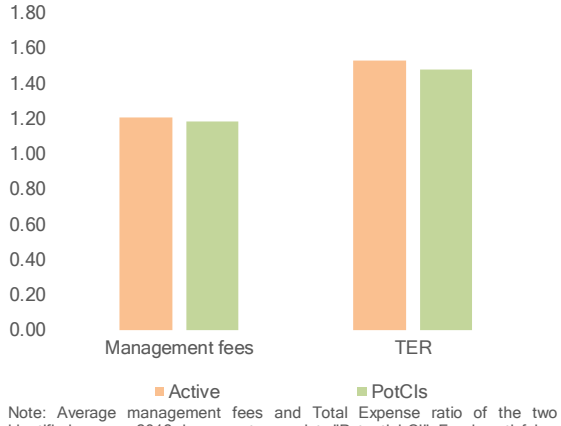


Note: "Active Share" =average value of fund AS in the sample. "Top mid-tail 15" =distribution between the 75th and 90th percentile. "Bottom mid-tail 15" =distribution between the 10th and 25th percentile.
Sources: Morningstar Direct, ESMA.

Chart 20

Average management fees and TER by *potCI* value

Little difference in management fees or TER

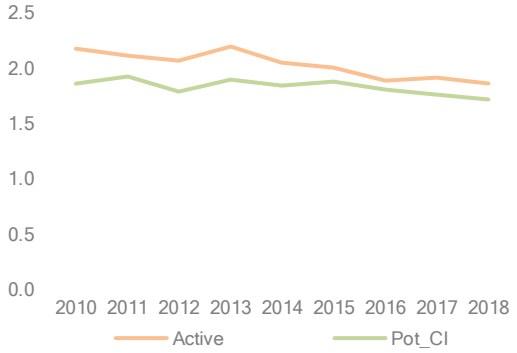


Note: Average management fees and Total Expense ratio of the two identified groups, 2018, in percentage points."Potential CI"=Funds satisfying the PotCI indicator variable in a given year, "Active"=other funds in the sample.
Sources: Refinitiv Lipper, Morningstar DIrect, ESMA.

Chart 21

Evolution of estimate total charges by *potCI* value

Similar pattern in costs if loads are included



Note: Estimated total charges, in percentage points, comprising TER and a component representingr front and back loads incurred by investors, weighted by annueal net flows. Methodology follows that in ESMA (2020). Total Expense ratio of the two identified groups, 2018, in percentage points."PotCI"=Funds satisfying the PotCI indicator variable in a given year, "Active"=other funds in the sample.
Sources: Refinitiv Lipper, Morningstar DIrect, ESMA.

Annex 2: Formal definitions

This Annex provides formal definitions of the closet indexing metrics outlined in section 2.1.

A2.1 Active Share

The AS of a fund captures the extent to which its portfolio differs from its benchmark index:

$$Active\ Share\ = \frac{1}{2}\sum_{i=1}^{N}\left|w_{fund,i} - w_{index,i}\right|$$

where $w_{fund,i}$ and $w_{index,i}$ are the portfolio weights of asset i in the fund and in the index respectively, and the sum is taken over all assets.[51] The absolute value of differences in these weights is taken so that over- and under-weightings in the fund relative to the benchmark count equally, while the normalisation factor of one half ensures that a fund that has no overlap with an index has an AS of 100%.

A2.2 Tracking Error

TE is defined as follows.[52]

$$Tracking\ Error\ = Stdev\left(R_{fund,t} - R_{index,t}\right)$$

where $R_{fund,t}$ and $R_{index,t}$ are respectively the returns of a fund and its benchmark index in period $t \in \{1, \dots, \tau\}$, and $Stdev(.)$ is the standard deviation taken over this period.

A2.3 Style-Shifting Activity

SSA is defined as follows.

$$SSA\ = \sum_{k=1}^{4}\left|b_{i,q,t}^{k} - b_{i,q-1,t}^{k}\right|$$

where $b_{i,q,t}^{k}$ is the kth factor from the 4-factor regression of the returns of fund i in excess of the risk-free rate for quarter q in half-year period t. SSA therefore measures quarterly changes in a fund's aggregate exposure to the Carhart factors of Box 2 in section 2.1.

[51] As in Cremers and Petajisto (2009), we restrict attention to equity funds and so the sum is taken over all equities held by funds or constituting part of a benchmark index.

[52] See e.g. Grinold and Kahn (1999).


esma
European Securities and
Markets Authority