The Global Rise of Index-Based Ownership of Firms

Finance Working Paper N° 1154/2026
August 2026

Bo Becker
Stockholm School of Economics, CEPR and ECGI

Rüdiger Fahlenbrach
Ecole Polytechnique Fédérale de Lausanne, Swiss Finance Institute, CEPR and ECGI

Ehsan Mahdikhani
Stockholm School of Economics

ECGI Working Paper Series in Finance

The Global Rise of Index-Based Ownership of Firms

Working Paper N° 1154/2026
August 2026

Bo Becker
Rüdiger Fahlenbrach
Ehsan Mahdikhani

We thank Patrik Adamson, Forrest Alogna, Marco Becht (discussant), Cees Bezuijen, Oscar Björklund, Patrick Bolton, Jill Fisch, Eduardo Geli, Mireia Giné, Jeff Gordon, Adam Green, Paul Huebner, Kenneth Khoo, Sebastian Lewis, Bernard Wang, Alexis Wegerich and Richard Winberg as well as seminar participants at SSE, SFI Lugano, GCGC 2026 and at the BI?

Abstract

We develop a simple and replicable five-variable model that identifies index-based investors with 98% accuracy. Our analysis shows that index ownership grew from 3.0% of global market capitalization in 2004 to 16.9% by 2023, driven by growth in explicit indexing. By the end of 2023, index-based investors own 10 percent or more of 5,300 firms globally. In the U.S., mid-caps have more index ownership than large-caps. In Asia and Europe, index ownership increases with firm size, and only the largest firms experience significant index ownership. Differences in market capitalization and free float help explain regional variations in index ownership.

Keywords: Passive Investing; Mutual Funds; Index Funds

JEL Classification: G11, G15, G32

Bo Becker

Stockholm School of Economics, CEPR and ECGI

Rüdiger Fahlenbrach

Ecole Polytechnique Fédérale de Lausanne, Swiss Finance Institute, CEPR and ECGI

Ehsan Mahdikhani

Stockholm School of Economics

The Global Rise of Index-Based Ownership of Firms*

Bo Becker† Rüdiger Fahlenbrach‡ Ehsan Mahdikhani§

August 31, 2026

Abstract

We develop a simple and replicable five-variable model that identifies index-based investors with 98% accuracy. Our analysis shows that index ownership grew from 3.0% of global market capitalization in 2004 to 16.9% by 2023, driven by growth in explicit indexing. By the end of 2023, index-based investors own 10 percent or more of 5,300 firms globally. In the U.S., mid-caps have more index ownership than large-caps. In Asia and Europe, index ownership increases with firm size, and only the largest firms experience significant index ownership. Differences in market capitalization and free float help explain regional variations in index ownership.

Keywords: Passive Investing; Mutual Funds; Index Funds

JEL Codes: G11, G15, G32

*We thank Patrik Adamson, Forrest Alogna, Marco Becht (discussant), Cees Bezuijen, Oscar Björklund, Patrick Bolton, Jill Fisch, Eduardo Geli, Mireia Giné, Jeff Gordon, Adam Green, Paul Huebner, Kenneth Khoo, Sebastian Lewis, Bernard Wang, Alexis Wegerich and Richard Winberg as well as seminar participants at SSE, SFI Lugano, GCGC 2026 and at the BI CCGR 2025 Conference for comments and suggestions. We thank Gaurav Yashvardhan and Hongyi Xu for excellent research support. Rüdiger Fahlenbrach thanks the Swiss Finance Institute for generous financial support. Bo Becker thanks the Nasdaq Nordic Foundation for a grant for this research.

†Stockholm School of Economics, CEPR and ECGI.

‡Ecole Polytechnique Fédérale de Lausanne, Swiss Finance Institute, CEPR and ECGI.

§Stockholm School of Economics.

1 Introduction

Index-based investment strategies have become a defining feature of modern equity markets. Their rapid and persistent expansion has given indexing a substantial footprint in the United States. As of year-end 2023, explicitly index-based domestic mutual funds and ETFs accounted for 18% of total U.S. stock market capitalization (Investment Company Institute, 2024). Using trading patterns around index deletions and additions in 2021, Chinco and Sammon (2024) estimate that the true number is twice as high once closet indexers and non-retail funds are included.

While aggregate holdings by index-based investors can be estimated with some confidence, identifying their ownership stakes in *individual* firms on a global scale is much more challenging. First, a large share of non-U.S. index investment is cross-border, so estimates based solely on domestic asset managers provide a poor approximation of total index ownership.[1] Second, asset management industries abroad are often less concentrated and less specialized, making it difficult to rely on a small set of large, index-focused firms such as BlackRock and Vanguard to infer investment styles. Third, the proliferation of national, regional, and global benchmarks complicates benchmark-based classifications of index investors (Cremers and Petajisto, 2009). Finally, data availability outside the United States has been typically more limited.

The rapid growth of indexing has generated substantial interest in its effects on firms and financial markets. Most evidence on the consequences of index ownership for firm behavior, corporate governance, liquidity, and price formation comes from the United States (e.g., Aghion et al., 2013; Boone and White, 2015; Appel et al., 2016; Schmidt and Fahlenbrach, 2017; Coles et al., 2022; Haddad et al., 2025; Kontz and Hanson, 2025).

[1] As an illustration, Cremers et al. (2016) convincingly demonstrate rising indexing in many mutual fund markets, but the ownership implications for individual firms are difficult to assess without knowing where index investment flows.

Applying this evidence to other markets requires understanding how prevalent index ownership is and which firms are exposed to it. This is particularly relevant because ownership is more concentrated in many countries, controlling shareholders are more prevalent, and legal, governance, and capital-market structures differ.

To study index ownership at the global level, we first need to measure it consistently at the individual firm level across countries. We introduce a new procedure that does this. We first classify a broad set of mutual funds, exchange traded funds (ETFs), and other investors as index or non-index based investors, using quarterly holdings data. We use the term index-based investors to describe investors whose portfolio holdings are chosen primarily to replicate the composition or returns of an externally defined index or rules-based benchmark, rather than through discretionary security selection.[2] We then aggregate index investor holdings at the portfolio firm level to produce a quarterly panel data of index ownership for public firms. Our procedure sidesteps any need for data such as third-party classifications of styles, maximizing the breadth and depth of our resulting data set.

We use our method on a global sample of public firms, spanning the period 2004-2023, and covering 76.2% of all shareholdings. Index ownership has increased from 3.0% of global market value in 2004Q1 to 16.9% of global market value by 2023Q4. Figure 1 shows the evolution over time. Several stylized facts are apparent. Index-based investors' ownership has grown rapidly and consistently. This expansion has substantially increased the number of firms for which index ownership is significant, shifting from a rarity twenty years ago to commonplace today. In 2004, 232 firms had index-based ownership of 10% or more. Only 19 of these firms were non-U.S. (firm geography not visible in the graph). Mid-sample (2013), the number of firms with index-based ownership exceeding 10% had risen to 2,416

[2]This definition includes full replication, sampling, optimized replication, and implicit benchmark tracking. It does not require the benchmark to be broad: a sector index fund is index-based if security selection follows a sector index. Leveraged and synthetic index products are index-based in their return objective, but enter our ownership measure only to the extent that they report direct equity holdings.

(a fifth were outside the U.S.). And at the end of our sample, over five thousand firms were above the 10% threshold, half of which were outside the U.S. These are the firms most likely to be affected by the governance, pricing, and liquidity consequences associated with high index ownership.

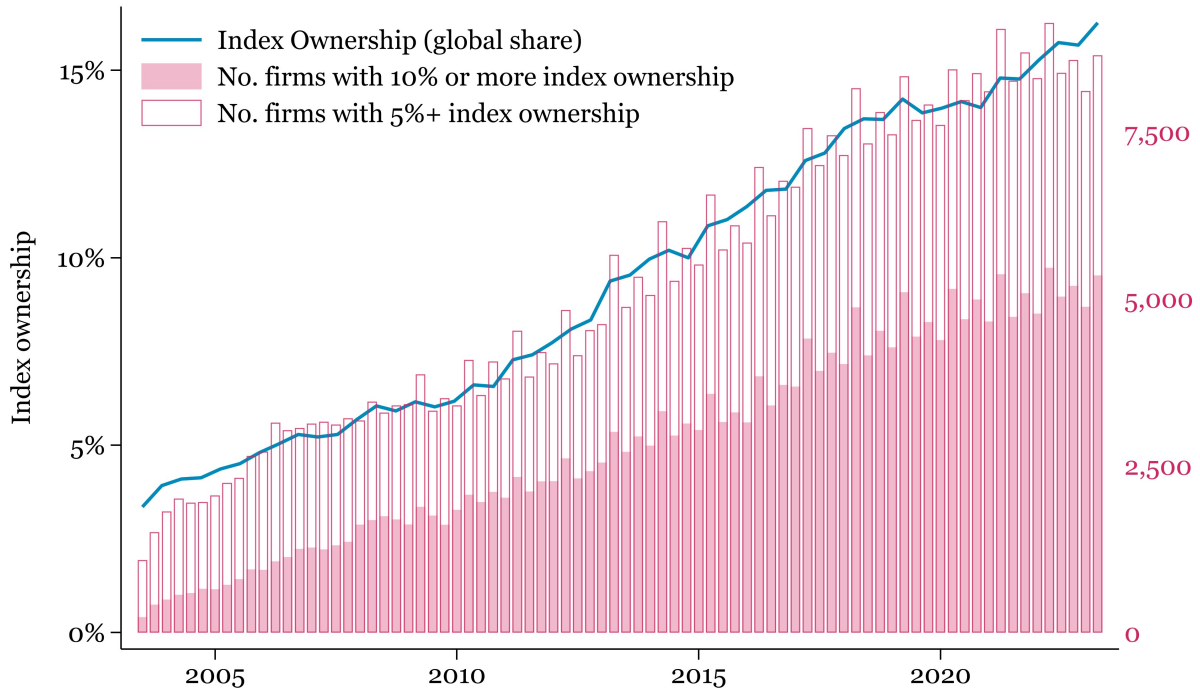

Figure 1: Index ownership. The number of firms globally with 5% or more (10% or more) index ownership is shown with pink bars (solid pink bars), against the right axis. The blue line shows total index holdings as a percentage of total market capitalization against the left axis.

The growth in indexing has been uneven across countries. In Europe, both Switzerland and the UK have passed 15% of market value, whereas other Western European countries have only passed 10% in the last few years, and in some cases experienced a plateau after Covid. In Asia, Japan has experienced a notable rise in index ownership, and Taiwan and Singapore are approaching 10%, South Korea and India now exceed 5%.

Second, there are important cross-sectional patterns. Index ownership is strongly related

to market capitalization everywhere, reflecting how indexes are designed. However, there are important regional differences in the index ownership-market cap relationship. In the U.S., the relationship is concave: medium-sized firms have the largest share of index ownership. In Europe and elsewhere, index ownership instead increases monotonically with firm size, and only the largest firms experience high index ownership. These patterns are not new phenomena, and are easily visible in earlier years, so they are not simply a question of market maturity. Rather, they may reflect underlying factors such as liquidity and family control, which limit the set of index-suitable firms.

Third, while a large number of indexes exists in all markets, the economically relevant benchmark universe differs sharply across regions. In the United States, index investing is highly concentrated in a small number of broad benchmarks (e.g., S&P 500, Russell 2000 and 3000, Nasdaq-100), so that a limited candidate set captures most passive investment. In contrast, outside the United States, index ownership is fragmented across national, regional, and global benchmarks, as well as provider-specific sector, style, and thematic index families. For a large European firm such as Siemens AG, inclusion spans multiple index families [3], together with numerous derived variants, implying hundreds of plausible benchmarks. As a result, approaches that classify indexers by matching portfolio holdings to benchmark compositions (e.g., Cremers and Petajisto, 2009) require the researcher to specify a benchmark set that is large, time-varying, and difficult to delineate consistently across markets.

A growing body of work measures passive or index ownership using fund-level holdings data, but typically either in aggregate form or within single-country settings. Appel et al. (2016) and Busse and Tong (2012) flag a fund as passively managed if the CRSP Mutual Fund Database classifies the fund as an index fund or if its fund name includes a string that identifies it as an index fund. Cremers and Petajisto (2009) classify funds with Active

[3] e.g., DAX, EURO STOXX, FTSE, STOXX Europe, STOXX Global, S&P Global, Dow Jones Global Titans, MSCI World, MSCI Europe, Low Carbon 100 Europe, and Dow Jones Sustainability

Share (half the sum of absolute differences between fund and benchmark weights) below 20% as passive or "pure indexing." Ferreira and Matos (2008) construct firm-level institutional ownership worldwide using Thomson Reuters data, showing that comprehensive cross-country holdings can be assembled but without distinguishing passive from active mandates. Cremers et al. (2016) combine Lipper fund data with FactSet/LionShares holdings across 32 countries. They use Active Share relative to benchmarks assigned by Lipper to distinguish closet indexers from truly active funds. Their approach demonstrates that benchmark-based classification can be implemented internationally when fund-level benchmark assignments are available. More recently, Antón et al. (2025) aggregate Thomson Reuters and FactSet holdings to estimate firm-level and sector-level common ownership in 49 countries, demonstrating the feasibility, but also the complexity, of harmonizing international ownership data. They focus on overlapping ownership, not on whether the investment style is based on indexing or not. U.S.-focused work by Chinco and Sammon (2024) goes further, by inferring implicit indexing from confidential fund-level regulatory data and reconstitution-day trading patterns, highlighting that benchmark-tracking ownership is underestimated in public filings. Collectively, these studies confirm that large-scale fund and ownership databases exist, yet they also reveal that accurately identifying *index* ownership at the firm level across countries remains challenging due to inconsistent fund classifications, varying disclosure regimes, and internal index mandates.

Industry sources such as the *Investment Company Institute Fact Book* (Investment Company Institute, 2024) and *the European Fund and Asset Management Association Fact Book* (European Fund and Asset Management Association, 2024) provide external validation that the global scale of indexed assets we infer is broadly consistent with industry aggregates, even though these practitioner datasets lack firm-level granularity, and may use different definitions of index-based investing. Consequently, while prior work demonstrates data feasibility, no study to date provides a harmonized, replicable, firm-level measure of index

ownership across countries of the kind we construct.

The rest of the paper is organized as follows. We introduce the data, motivate our benchmark-free approach, and develop the measurement algorithm in Section 2. We present empirical findings in Section 3. We conclude in Section 4 by discussing what the international evidence implies for the predominantly U.S.-based literature on index ownership.

2 Data and measurement

2.1 Data sets used

We use a large data set of portfolio holdings of ETFs, mutual funds and other investors from S&P Global Market Intelligence, which reports to cover ownership data for 99.7% of public companies globally. The data is called "Ownership, Transactions and Key Developments" and is based on processing primary sources such as regulatory filings (e.g., in the U.S., forms 13F, 13D, 13G, 3, 4, and 5 filed with the Securities and Exchange Commission (SEC); in Canada, filings through the System for Electronic Document Analysis and Retrieval (SEDAR) administered by the Canadian Securities Administrators (CSA)); company reports (Annual reports, proxy statements such as DEF 14A and press releases); open- and closed-end fund and ETF reports on holdings. The data covers more than 36,000 institutional investment firms for the period 2004-2023.

When we have more than one portfolio snapshot in a quarter for a given fund, we use the observation with the latest reporting date. We combine the portfolio data with variables describing firm characteristics, such as market capitalization, from S&P Capital IQ. We measure free float using S&P Capital IQ's variable IQ_FLOAT_PERCENT, defined as the percentage of common shares outstanding that are freely available for public trading. S&P Capital IQ excludes insider holdings and strategic 5% shareholders. We use Morningstar's

classification of fund strategies to calibrate our model for identifying index investors.

2.2 Inferring index investors from benchmark matching

Identifying index-based investors requires distinguishing portfolios that track an index from portfolios that are actively managed. One approach is to compare each portfolio to the holdings or returns of candidate benchmark indexes. Cremers and Petajisto (2009), for example, define Active Share as one half of the sum of absolute differences between a fund's portfolio weights and its benchmark weights. Because the true benchmark is often unobserved, they compute Active Share relative to a set of candidate benchmarks and assign each fund to the index that minimizes Active Share. In their U.S. application, this benchmark set consists of nineteen broad indexes from S&P/Barra, Russell, and Wilshire. Petajisto (2013) extends this approach by combining Active Share with tracking error to distinguish stock pickers, factor bettors, closet indexers, and index funds.

Benchmark-based approaches can be extended internationally when reliable fund-level benchmark assignments are available. Cremers et al. (2016), for example, use benchmarks independently assigned by Lipper to calculate Active Share for mutual funds across 32 countries. They restrict the benchmark universe to 88 economically important world, regional, and country indexes and construct benchmark portfolio weights from the holdings of explicitly indexed funds tracking each benchmark. This approach is well suited to their analysis of competition in the mutual fund industry. Our objective requires broader coverage. To measure index-based ownership at the firm level, a benchmark-based approach would require benchmark assignments for a much wider set of investors and sufficiently comprehensive benchmark coverage that omitted indexes account for little index-based ownership. These requirements are difficult to satisfy in a global ownership sample for three reasons.

First, the relevant benchmark universe is difficult to delineate. In the United States,

a small set of broad indexes captures much of indexed equity investment. Outside the United States, index investment is spread across national, regional, and global benchmarks, and across provider-specific sector, style, and thematic index families. The set of potentially relevant benchmarks is therefore large, time-varying, and partly unobservable.

Second, the classification is sensitive to the benchmark set. Active Share and tracking error depend on both the investor's portfolio and the benchmark used for comparison (e.g., Frazzini et al., 2016). This creates a tradeoff in benchmark selection. A narrow benchmark set can misclassify funds tracking omitted indexes as active. An overly broad set increases the likelihood that some benchmark fits an active portfolio closely. In a global sample, the benchmark-selection tradeoff varies across markets because benchmark coverage is more complete in some countries and regions than in others.

Third, tracking-error-based classification requires realized portfolio returns (e.g., Petajisto, 2013). These returns are observed in some mutual fund return databases but not for all vehicles in a global holdings panel. Reconstructing them from quarterly holdings would require assumptions about unobserved portfolio changes and positions. Tracking-error classification also retains the benchmark-selection problem.

We therefore use a benchmark-free classification. Rather than asking which benchmark a portfolio tracks, we ask whether the portfolio exhibits characteristics of index-based investment. The next two sections describe the theoretical motivation and implementation of our classification method.

2.3 Inferring index investors from portfolio holdings

In the simplest case, identifying index-based investors from portfolio holdings is straightforward.[4] Let B denote a value-weighted benchmark consisting of a set of firms, let M_i denote the market capitalization of firm i, and let the total benchmark market

[4]For notational simplicity, we suppress time subscripts throughout this subsection.

capitalization be

$$M_B = \sum_{j \in B} M_j. \tag{1}$$

A fund f with assets under management A_f that perfectly tracks the value-weighted benchmark B holds H_{fi} dollars in firm i:

$$H_{fi} = A_f \frac{M_i}{M_B}, \qquad i \in B. \tag{2}$$

Its ownership share in firm i is therefore

$$o_{fi} = \frac{H_{fi}}{M_i} = \frac{A_f}{M_B}, \tag{3}$$

which is constant across all benchmark constituents. In this idealized setting, identifying an index fund reduces to testing whether ownership shares are constant across benchmark constituents.[5]

Today, benchmarks are typically based on free-float-adjusted market capitalization rather than total market capitalization. Let ϕ_i denote the free-float factor for firm i used by the provider of benchmark B.

Let

$$M_B^{FF} = \sum_{j \in B} \phi_j M_j \tag{4}$$

denote the free-float-adjusted market capitalization of the benchmark.

A fund f with assets under management A_f that perfectly tracks a free-float-weighted

[5]Equal-weighted benchmarks imply a different restriction. If benchmark B has N_B constituents and fund f tracks it exactly, then $H_{fi}/A_f = 1/N_B$ for all $i \in B$. Thus, portfolio weights are constant across benchmark constituents, not ownership shares. Our derivation focuses on value-weighted and free-float-weighted benchmarks, because they are the dominant index type.

benchmark holds H_{fi} dollars in firm i:

$$H_{fi} = A_f \frac{\phi_i M_i}{M_B^{FF}}, \qquad i \in B. \tag{5}$$

Its free-float-adjusted ownership share is therefore

$$s_{fi} \equiv \frac{H_{fi}}{\phi_i M_i} = \frac{A_f}{M_B^{FF}} \equiv c_f, \tag{6}$$

which is again a constant, c_f, across benchmark constituents.

However, the derivation above assumes that the researcher observes the same free-float adjusted market capitalization used by the benchmark provider to construct benchmark weights. In reality, different benchmark providers employ different free-float adjustments, which would require researchers to know which benchmark family is being tracked. In addition, benchmark providers often use proprietary free-float adjustments, including investability screens. Hence, a researcher using portfolio holdings of a fund and a generic free-float adjustment from a global data provider will almost certainly measure ϕ_i with error, and observe non-constant s_{fi} even for an index fund that perfectly replicates its benchmark.

Let ϕ_i denote the benchmark provider's unobserved free-float factor and let $\hat{\phi}_i$ denote the free-float factor used by the researcher. The researcher calculates

$$\hat{s}_{fi} = \frac{H_{fi}}{\hat{\phi}_i M_i}, \tag{7}$$

and $\hat{s}_{fi}$ need not equal c_f because $\hat{\phi}_i \neq \phi_i$. Let

$$\hat{s}_{fi} = c_f + m_{fi}, \tag{8}$$

where m_{fi} captures deviations from the constant ownership share of fund f in firm i arising from imperfect observation of the benchmark provider's methodology.

The derivation above also assumes full replication of benchmark weights. In practice, benchmark-tracking funds seek to replicate benchmark returns rather than benchmark holdings. Consequently, they may hold only a subset of benchmark constituents or deviate from constant ownership shares for implementation reasons. Let η_{fi} denote the resulting deviation in fund f's free-float-adjusted ownership share in firm i.

Then

$$\hat{s}_{fi} = c_f + m_{fi} + \eta_{fi}. \tag{9}$$

Relatedly, the active-share literature emphasizes that many funds occupy an intermediate position between pure indexing and fully active management (Cremers and Petajisto, 2009; Petajisto, 2013). Such funds may remain close to a benchmark while deviating from benchmark weights in individual securities. In our framework, these deviations are reflected in η_{fi}. From the perspective of measuring the aggregate importance of index investing, these "closet indexers" are often more similar to benchmark-tracking investors than to actively managed funds. Our classification approach is designed to accommodate such indirect benchmarking without knowing the benchmark.

In practice, therefore, even purely benchmark-tracking portfolios need not display constant free-float-adjusted ownership shares because the researcher does not observe the precise benchmark universe, the benchmark provider's proprietary float adjustments, or the investor's implementation choices. The empirical problem is to identify portfolios whose holdings are systematically close to the behavior implied by index-based investment. Our classification uses portfolio-level characteristics that capture the key features of indexes without knowing anything about the followed benchmarks. Index-based investors should trade relatively little, hold many positions, have relatively similar free-float-adjusted ownership shares across portfolio firms, and make position changes that are dispersed less

like stock-specific bets and more like proportional rebalancing. The next section describes how we construct such variables and combine them into a benchmark-free measure of index-based ownership.

2.4 Classification method

We classify funds using four portfolio characteristics: Variance, Position Changes, Turnover, and Portfolio Holdings. The first three are motivated directly by the benchmark-tracking framework developed in the previous subsection, while the last captures the breadth of benchmark exposure. We calibrate the methodology against Morningstar's classification of fund types.[6]

The first variable is **Variance**, the within-portfolio dispersion of ownership shares across securities. We begin with the free-float-adjusted ownership share of each position, scale it by the portfolio's average ownership share, truncate the resulting measure at the 10th and 90th percentiles across securities, and then take the standard deviation across positions divided by their mean. As discussed in the previous subsection, benchmark tracking implies approximately proportional ownership shares across securities. **Variance** therefore measures the extent to which observed ownership shares deviate from this proportionality restriction. We expect index investors to have low values of **Variance**.[7]

The second variable is **Position Changes**. For each security held by a fund, we calculate the percentage change in its free-float-adjusted ownership share relative to the previous reporting period. We then compute the standard deviation of these percentage changes divided by their mean absolute value. If a fund increases or decreases all positions proportionally, the variable equals zero. As discussed above, benchmark tracking implies

[6]We can also calculate index ownership directly using Morningstar's classification. However, this approach would omit many non-U.S. funds that are not classified by Morningstar, as well as closet indexers.

[7]We use the free-float-adjusted version of Variance (i.e., using shares owned divided by free-float adjusted shares outstanding) because it better matches the methodology of most index families, but obtain very similar results using unadjusted ownership shares (i.e., shares owned divided by total shares outstanding).

relatively stable ownership structures over time. **Position Changes** therefore captures the extent to which changes in ownership shares depart from proportional rebalancing. We expect index investors to have low values of **Position Changes**.

The third variable is **Turnover**, measured as the ratio of purchases plus sales to portfolio value. In the framework developed above, large changes in ownership shares correspond to larger implementation deviations. We therefore expect benchmark-tracking funds to trade less frequently than active managers and to exhibit lower values of **Turnover**.

The fourth variable is **Portfolio Holdings**, measured as the number of securities held by the fund. Unlike the previous three measures, this variable does not arise directly from the proportionality restriction. Instead, it captures the breadth of benchmark exposure. Benchmark investors typically seek diversified market exposure and therefore hold a larger number of securities than active managers.[8] We therefore expect index investors to have high values of **Portfolio Holdings**. Detailed definitions of all variables are provided in Appendix A, Table A1.

Each of the four quantitative variables exhibits a markedly different distribution for funds classified as index and non-index by Morningstar. We standardize each variable by subtracting the mean and dividing by the standard deviation of the non-index fund distribution, so that zero corresponds to the average non-index fund. Figure 2 shows that relative to this benchmark, index funds exhibit lower turnover, greater breadth of holdings, lower dispersion in free-float-adjusted portfolio weights, and less cross-sectional variation in position adjustments across holdings (i.e., more proportional rebalancing). The standardized distributions for index funds are therefore systematically shifted away from the non-index benchmark across all four dimensions. Together, these patterns indicate that the four variables sharply distinguish index from non-index investment behavior.

[8]Portfolio Holdings will be less predictive for funds tracking some older national indexes outside the United States. For example, the Swiss Market Index (SMI) contains 20 stocks, while the German DAX and French CAC 40 each contained 40 constituents in 2023.

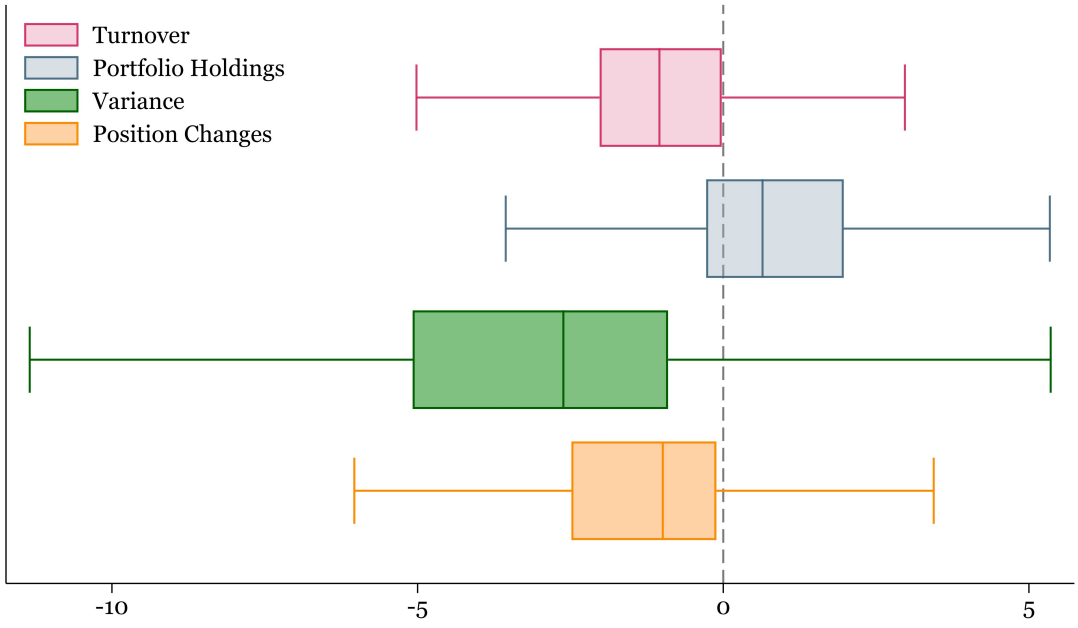

Figure 2: Characteristics of index funds relative to non-index funds. The figure shows the distribution of four variables in the subsample of funds classified as index funds by Morningstar. The distribution of each variable has been normalized by the mean and standard deviation of the distribution for *non-index* funds (therefore, units on the x-axis are standard deviations, based on the non-index population of funds). For each variable, the boxes show the median, 25th and 75th percentiles of the normalized distribution, and the whiskers show 1.5 times the interquartile range on each side.

Our final variable is an indicator variable equal to one if the fund's business description or name suggests that the fund is an index fund, and zero otherwise. For approximately 32% of all funds in our database, S&P Capital IQ provides a business description (field *IQ_BUSINESS_DESCRIPTION*). Appendix B contains three examples of these descriptions. We use a machine learning algorithm that assesses whether the business description mentions tracking an index rather than merely benchmarking the fund's performance to an index, or whether it mentions replication techniques such as full replication or representative sampling. For those funds where we lack a business description, we assign a value of one to the indicator variable if the fund's name contains the name of a benchmark

or index family and zero otherwise.[9] Appel et al. (2016) and Busse and Tong (2012) have used similar name-based classifications for a U.S. sample of mutual funds.

Unlike the previous four variables, this indicator variable does not capture portfolio behavior. Instead, it captures explicit index affiliation. At the same time, a business description and name-based approach alone would fail to identify closet indexers and other undeclared benchmark followers. We therefore view the indicator variable as a complementary signal alongside the portfolio-based measures developed above.

To identify index funds, we first estimate a model for the set of fund-quarters that are classified by Morningstar. The regression combines the four quantitative variables in a single probit regression:

$$Pr(I_{it}^{\text{Index Fund, Morningstar}} | \mathrm{X}_{it}) = \Phi\Big(\alpha + \theta\,\text{Turnover}_{it} + \gamma\,\text{Portfolio Holdings}_{it} + \mu\,\text{Variance}_{it} + \lambda\,\text{Position Changes}_{it}\Big) \quad (10)$$

The dependent variable is an indicator equal to one if the fund is classified as an index fund by Morningstar, and Φ denotes the standard normal CDF.

We then apply the estimated coefficients to fund-quarters without Morningstar classifications (about two thirds of the underlying fund-quarter data) and classify a fund-quarter as indexed if its fitted probability exceeds the classification threshold $p^* = 0.5$.[10]

[9] We use the following list of terms: Index, CAC, DAX, OMX, FTSE, S&P, S&P 500, SP500, MSCI, Nasdaq, Russell, Total Market, QQQ, SPY, IVV, VOO, VTI, IWM, Set 50, Jumbo 50, Jumbo 25, Select Sector SPDR, Euro Stoxx, Topix, Nikkei, JASDAQ, IBEX, TSX, ASX, Bovespa, KOSPI, Hang Seng, HSI, Lyxor Index, Dow, iShares Core, Vanguard Index, Vanguard 500, Vanguard Total Stock, Xtrackers, Schwab 1000, Fidelity 500.

[10] With $p^* = 0.5$, classification as an index fund requires the fitted latent probit index to be non-negative; because the estimated constant is -0.09, this is equivalent to imposing a cutoff of 0.09 on the linear combination of the four fund characteristics.

$$I^{\text{Index Fund Probit}} = \begin{cases} 1 & \text{if } -0.19 \times \text{Turnover} + 0.05 \times \text{Portfolio Holdings} \\ & \quad - 0.06 \times \text{Variance} - 0.04 \times \text{Position Changes} \geq 0.09, \\ 0 & \text{otherwise} \end{cases}$$

The estimate is based only on features of the portfolio, but still works very well: the model replicates the Morningstar classification for 86% of fund-quarters, and 94% of AUM.

In robustness tests, we choose p^* to deliver a precision of 90% in the Morningstar-classified sample, where precision is defined as

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}.$$

Thus, $p^* = p^{\text{prec}}$ is the lowest cutoff for which at least 90 percent of fund-quarters classified as index funds are also classified as index funds by Morningstar. Using p^{prec}, we classify 93% of fund-quarters and 97% of AUM correctly.

We can improve the fit by including information in fund business descriptions and names. To incorporate that information, we combine the predicted probit variable with the business-description/name indicator as follows: $I^{\text{Index Fund}} = \max\{I^{\text{Index Fund Probit}}, I^{\text{Index BD or Name}}\}$. If only one of the arguments is defined, we use that element. The resulting classification is available for each fund and quarter.[11] The advantage of using a max procedure, rather than simply including the business description/name indicator variable in the probit model, is that we might better capture closet indexers (those funds that exhibit a portfolio close to an index without explicit index affiliation in either the

[11]The business-description/name indicator is time-invariant. The portfolio-based classification varies by quarter as fund holdings change. In practice, there is little variation in the resulting classification within funds over time. As an alternative, we classify a fund as an index investor in all periods if it is classified as such in any period. Results are qualitatively and quantitatively similar. We prefer the period-by-period classification because it allows the holdings-based signal to vary over time.

business description or fund name). Our broader model agrees with Morningstar's classification for 95.7% of fund-quarters and 98.4% of AUM. Disagreements are about equally common in either direction.[12] The excellent fit means that we can accurately predict index status using the four portfolio characteristics plus the business-description/fund name information. An important caveat is that we use fund descriptions as of 2026 when we downloaded the data – there is no panel of older descriptions.

Based on the classification of funds as index or not and each fund's ownership share of publicly listed companies, we can calculate index ownership for each company and each quarter by aggregating the ownership across all index funds and index ETFs at the end of each quarter. The procedure leads to an unbalanced panel, with a quarterly observation for each quarter that a company is public, within the 2004-2023 period. There are 2.2 million observations, 80 quarters, and 78,387 firms in total. Of those, 10,251 firms are present during the whole sample period. In the last cross-section (2023Q4), there are 31,892 firms.

In the whole sample, average index ownership is 3.3%, and median is 0.3%.[13] The average rises from 1.3% at the start of the sample to 4.9% at the end. We divide index ownership into explicit and closet indexing. Explicit indexing comprises ownership by funds and ETFs whose business descriptions indicate index tracking or replication or, when business descriptions are unavailable, whose names refer to indexing or include an index name. Closet indexing comprises all remaining ownership classified as index-based by our algorithm. Average and median closet index ownership is 1.6% and 0.12%, and average and median explicit indexing is 1.7% and 0.01%.

[12]Two examples of funds where our algorithm classifies the fund as an index fund, but Morningstar does not, include the "Horizon Defensive Core Investor" which is classified as "Large Blend" by Morningstar. This fund had 70-90 holdings in 2020 and 2021, but then gradually increased this to over 300 in late 2023. Another example is the "SEI Dynamic Asset Allocation" fund, whose largest holding in October 2025 was a 17.2% allocation to E-mini S&P500 futures. Plausibly, our algorithm provides a justifiable categorization in both cases.

[13]Index ownership is strongly increasing in market capitalization, so value weighted numbers are considerably higher than equal weighted numbers. For example, the value weighted average index ownership in the whole sample is 11.0%, three times the equal weighted average.

Finally, for benchmarking purposes, we also calculate insider ownership for each firm in each quarter, based on S&P's classification (insider ownership is the firm-quarter percentage of shares outstanding held by all individual holders that S&P Global Market Intelligence flags as corporate insiders, i.e., officers, directors, and other beneficial owners per regulatory filings; we aggregate across all such individuals). Average insider ownership is 12.3% of shares, and the median is 2.2%.

	Fund-quarters classified (%)	Fund-quarters classified (%)	Agreement with Morningstar	Agreement with Morningstar (AUM)
Sample	World	ex-U.S.	World	World
CRSP MF database	634,251 (29.6%)	0 (0%)	n/a	n/a
Morningstar	1,120,978 (52.3%)	638,007 (44.7%)	n/a	n/a
Full probit model	2,142,814 (100%)	1,428,656 (100%)	95.7%	98.4%

Table 1: Comparison of index fund classifications. Summary of alternative methods of identifying index funds. Row 1 refers to the Center for Research in Security Prices Mutual Fund Summary database, available from WRDS (CRSP coverage is annual). Row 2 refers to Morningstar's index flag, used in this paper. Row 3 refers to the model used in this paper based on S&P Global Market Intelligence (see Section 2.4). Numbers in parentheses show which fraction of our database can be classified. AUM refers to weighting by the aggregate value of securities in each fund's portfolio.

How well does our classification perform compared to alternatives? In Table 1, we compare three classification approaches: the Morningstar classification of funds, the CRSP Mutual Fund database index fund flag, and our own model which combines data from fund names and descriptions with a probit model of portfolio characteristics. For each approach, we report the number of fund-quarters available during our sample period, globally and ex-U.S. (since we need a measure of index ownership that is viable outside the U.S.). We also report the agreement with the Morningstar index identifier, which we treat as the target variable to match for our own model. Our model has several attractive features: it is available for a large number of U.S. fund-quarters, has great coverage also of non-U.S. fund-quarters, and achieves excellent alignment with Morningstar in the sub-sample

where Morningstar's index classification is available. The value-weighted agreement is particularly high (98.4%), which is key for understanding ownership patterns (since large funds own more shares, they matter more to aggregate index ownership). The combination of high accuracy and broad coverage makes our classification method particularly useful for assessing the role of index ownership globally.

3 Empirical patterns in index ownership

Our firm-level panel data set allows us to examine variation in index ownership across countries, over time, and across firms. We first document cross-country differences in index ownership, then examine the evolution of explicit and closet index ownership alongside insider ownership, and finally study cross-firm variation in index ownership and index suitability. We leave several questions for future work, including a comprehensive analysis of the global firm-level consequences of indexing.

3.1 Cross-country patterns

The overall growth in index investing, as seen in Figure 1, is rapid and smooth, but it hides considerable variation across firms and countries. Figure 3 compares the growth in indexing in the U.S. to that in other markets. In Panel A, we compare selected Asian countries to the U.S.: whereas China has seen little or no growth in indexing since around 2010, India, South Korea and Singapore show steady growth from low levels, Taiwan experienced somewhat faster growth, and Japan shows accelerating growth, narrowing the gap with the U.S. Japanese firms have the second highest index ownership after U.S. firms.

Panel B compares European countries to the U.S. Both Switzerland and the U.K. have seen rapid growth, trailing the levels of U.S. index ownership by around ten years. The Netherlands and the Nordics have experienced somewhat slower index ownership growth,

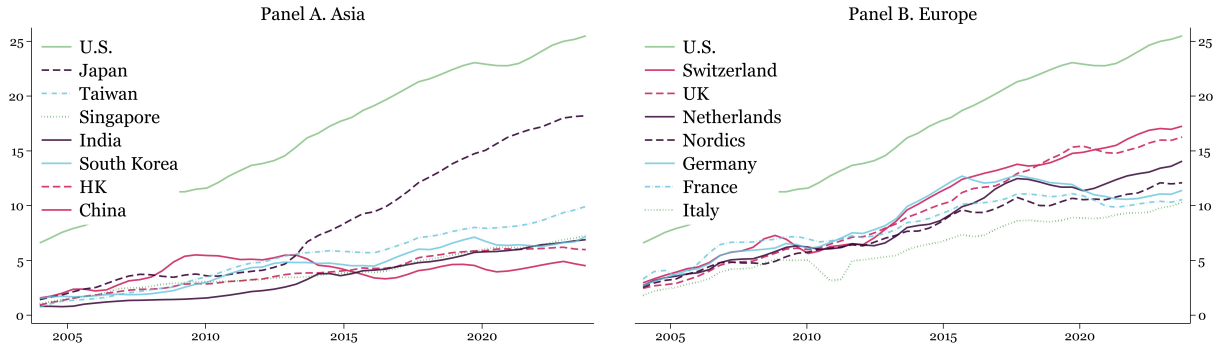

Figure 3: Index ownership of firms in selected markets. This figure plots the value-weighted index ownership for the U.S. and selected Asian (Panel A) and European (Panel B) countries, 2004-2023.

and Germany, France and Italy have seen the slowest growth. Note that the lines correspond to the index ownership shares of *firms* in each country, and not funds. For many European firms, much of the index ownership stems from indexes at the global or European, not national level, which is suggestive of an investor base from abroad (e.g., U.S. or other European investors). Because of these large cross-country differences, the implications of index ownership for liquidity, market efficiency, or corporate governance may also differ across countries.

3.2 Explicit and closet index ownership and the role of insiders

Closet index funds, i.e., funds whose portfolios exhibit characteristics of index investing without explicitly reporting it, tend to charge higher fees and underperform (Cremers et al., 2016; Bjerksund et al., 2026). We estimate closet indexing as the difference between total index ownership (identified by our algorithm) and ownership by explicit indexers, identified from business descriptions that indicate index tracking or replication or, when business descriptions are unavailable, from fund names that refer to indexing or include an index name.

In Figure 4, we plot the evolution of value-weighted explicit and closet index ownership for the world and for large regions. Several patterns are noteworthy. The relative importance of explicit and closet indexing differs substantially across regions. In the U.S., the growth in index ownership is driven almost entirely by explicit indexing, while closet index ownership has changed relatively little since around 2010. In Europe and Asia, by contrast, explicit and closet index ownership have both increased substantially and by similar magnitudes. Explicit indexing has only recently exceeded closet indexing in these regions. The worldwide pattern, in which explicit indexing accounts for most of the growth in index ownership, therefore largely reflects the U.S. experience.

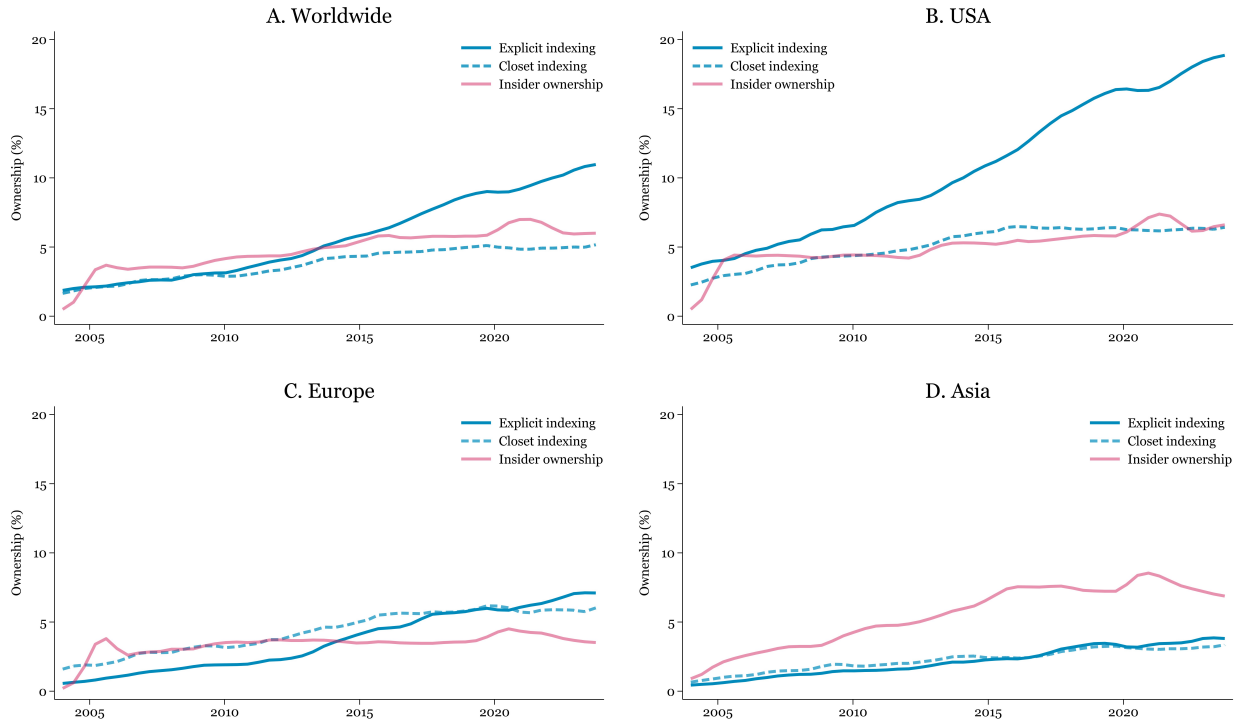

Figure 4: Explicit indexing, closet indexing, and insider ownership. This figure plots the evolution of explicit indexing, closet indexing, and insider ownership. Panel A shows data for the world, Panel B U.S., Panel C Europe, and Panel D Asia.

Figure 4 also reports insider ownership, a natural candidate for explaining regional

differences in index ownership. Insiders and other long-term shareholders, such as corporations and strategic institutions, are often excluded from index float adjustments. For example, MSCI excludes holdings that are unlikely to be available in the market when calculating free-float-adjusted market capitalization, and S&P makes similar adjustments. Higher insider ownership therefore reduces the shares available for index investment.

The aggregate patterns in Figure 4 suggest, however, that insider ownership alone cannot explain the regional differences in index ownership. Europe has substantially less index ownership than the U.S. despite broadly similar levels of insider ownership. Insider ownership is higher in Asia, consistent with lower free float contributing to lower index ownership, but the difference is small relative to the much larger gap in index ownership. However, these aggregate patterns do not rule out an important role for free float at the firm level or for other firm characteristics in explaining index ownership. We therefore turn to firm-level variation in index suitability.

3.3 Cross-firm patterns

Cross-firm variation in index ownership is closely related to how broadly index families cover the market. The U.S., with many indexes such as the Russell 3000 comprising thousands of firms including small-caps, and even dedicated indexes for mid-caps and small-caps such as the S&P MidCap 400 and S&P SmallCap 600 Index, is internationally an outlier. Long-standing national indexes in many countries focus on a small set of the largest domestic firms (e.g., the DAX, which includes forty large German firms; the CAC 40, which includes forty large French firms; or Singapore's Straits Times Index, which includes thirty large firms). Broader non-U.S. indexes with more stocks tend to be cross-country but still focus on large-cap and mid-cap stocks (e.g., the MSCI Europe has around four hundred constituents from fifteen countries). Liquidity and free-float considerations may help explain the limited representation of small-cap stocks in non-U.S. indexes. Most, but

not all, indexes are value-weighted, and many use free-float adjustments.

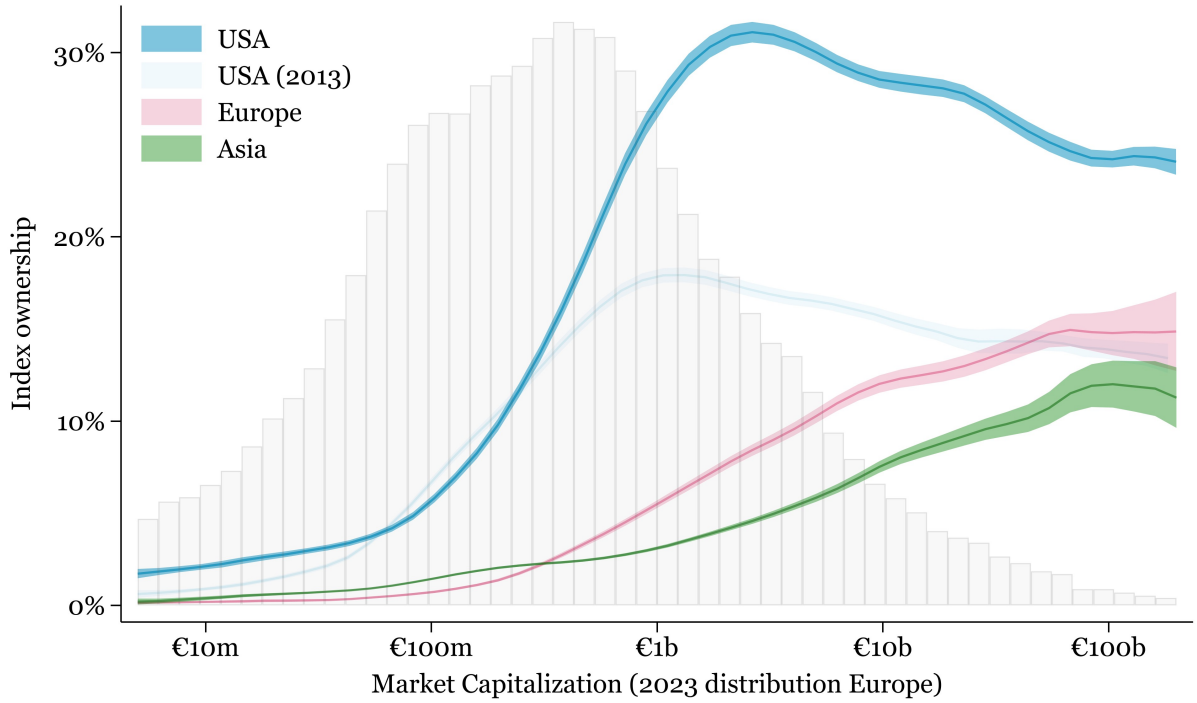

Figure 5: Index ownership and firm size. The figure plots the cross-sectional relationship between firm market capitalization and index ownership in the U.S., Asia and Europe in 2023, as well as for the U.S. in 2013. Market capitalization is in real values (2023 Euros). The distribution of market capitalization is indicated in the background, based on the 2023 subsample.

These differences in index coverage suggest that the relation between firm characteristics and index ownership may differ across markets. We first examine market capitalization. Figure 5 summarizes the relationship between firm size and index ownership. We focus on 2023, but as a benchmark, the figure also shows the U.S. relationship for 2013; the regional patterns persist in earlier years, at lower levels. In the U.S., index ownership is low for small-caps and rises rapidly for mid-caps once market capitalization exceeds a few hundred million euros. Index ownership peaks at around 30% for firms worth a few billion euros. Beyond that point there is a decline, and the largest firms have less index ownership than mid-caps. Rowley et al. (2025) document a similar pattern in their Figure 5

for stocks in the Russell 3000, with index-fund ownership peaking in the mid-cap range and declining for the largest firms.

Outside the U.S., on the other hand, the relationship between market capitalization and index ownership is monotonically positive. Small-caps have low or no index ownership. Mid-caps have more, but much less than comparable U.S. firms, about a quarter of the U.S. level in Europe and less in Asia. Index ownership then increases in the large-cap segment, to about half of the U.S. level in Europe and slightly less in Asia. The broader coverage of U.S. index families is consistent with these differences. Well-established index families such as S&P and Russell cover the mid- and small-cap segments extensively, while many non-U.S. indexes remain concentrated among larger firms. Greater liquidity and free float may also facilitate index inclusion further down the market-cap distribution in the U.S.

If the low index ownership of non-U.S. mid-caps reflects limits to their indexability, the aggregate capacity for index ownership may be lower outside the United States. To the extent that index ownership benefits firms and U.S. listing facilitates inclusion in U.S. indexes, the higher index ownership of U.S. mid-caps may also increase the attractiveness of a U.S. listing for firms in this size range.

Figure 5 shows a strong relation between market capitalization and index ownership, but market capitalization is only one dimension of index suitability. We therefore examine the firm-level correlates of index ownership in a multivariate regression. In addition to market capitalization, we include free float, which measures the supply of shares available to public-market investors, and trading volume, which captures trading capacity. A firm may be large and have substantial free float but still be costly to include in an index if little of its equity trades. We also include dividend payer status and positive profitability, which capture simple rules-based screens relevant for dividend- and quality-oriented index products.

Table 2 examines the predictive variables jointly in regressions of quarterly firm-level

	(1)	(2)	(3)
	Index Ownership		
Market Cap (log)	**1.114*****	**0.838*****	**0.869*****
	(0.259)	*(0.149)*	*(0.143)*
Free float	**0.103*****	**0.065*****	**0.066*****
	(0.034)	*(0.018)*	*(0.019)*
Trading Volume (log)	**0.192**	**0.371***	**0.334****
	(0.361)	*(0.191)*	*(0.167)*
Dividend payer	**-0.267**	**0.748****	**0.679****
	(0.298)	*(0.377)*	*(0.335)*
Positive earnings	**-0.032**	**0.542**	**0.586**
	(0.142)	*(0.374)*	*(0.440)*
R-squared	0.338	0.545	0.589
Observations	1.756m	1.756m	1.755m
Country FE	-	Yes	-
Time FE	-	Yes	-
Country x Time FE	-	-	Yes
Error clusters	Country	Country	Country

Table 2: Firm-level variation in index suitability. The table reports results from a regression of quarterly firm-level index ownership on a free float measure, the logarithm of market capitalization (measured in 2023 Euros), the logarithm of the average annual daily trading volume (measured in millions of 2023 Euros), an indicator variable equal to one if the company paid dividends and zero otherwise, and an indicator variable equal to one if earnings from continuing operations were positive, and zero otherwise. Column 2 includes time- and country-fixed effects, and column 3 contains country $\times$ year fixed effects. The letter m refers to millions. Coefficients are reported in bold font, and robust standard errors clustered by country are reported below. *, **, and *** denote statistical significance at the ten, five, and one percent levels, respectively.

index ownership on firm characteristics. The first regression omits country fixed effects (and is therefore partially identified by cross-country variation). Market capitalization is a strong predictor of index ownership in all three specifications, with a coefficient of about one. This implies that every order of magnitude increase in market capitalization predicts about 2 percentage points more index ownership. Because the average log market capitalization is higher in the U.S., this can explain part of the regional gap in index ownership compared to Europe and Asia. Using the coefficient in column three, about one-eighth of the index ownership gaps (in the full sample) can be explained by market cap. Free float is also associated with higher index ownership in all three specifications. Every 15 percentage points of free float increases index ownership by around one percentage point in the fixed-effect regressions. The value-weighted average of free float in the U.S. is 91%, in Europe 74% and in Asia 53%. These differences explain about one-tenth of the index ownership gap with the U.S.

The coefficient on the dividend payer indicator is significant in the fixed effect regressions, but insignificant in the first column (perhaps reflecting that U.S. firms often do not pay dividends but have high index ownership). Relative to the free float and market cap variables, the economic magnitude is modest: for a value of one for dividend payer (vs. zero), index ownership is predicted to increase by 0.7%. Trading volume is positively associated with index ownership in the fixed-effect specifications, although the economic magnitude is modest. In column 3, a doubling of trading volume is associated with about 0.23 percentage points more index ownership. Whether a company is profitable or not does not predict index ownership.

Overall, the cross-section of index ownership is strongly related to size and free float. Index investors hold larger stakes in firms that are large and investable, while trading volume, dividend status, and profitability add little incremental explanatory power. These patterns reflect how index providers design benchmarks, which benchmarks asset managers choose

to track, and how investors allocate assets across funds.

4 Conclusion

Prior research links index ownership to firm behavior, corporate governance, liquidity, and price formation (Appel et al., 2016; Boone and White, 2015; Schmidt and Fahlenbrach, 2017; Haddad et al., 2025). Most of this evidence comes from the United States, while much less is known about the prevalence, distribution and implications of index ownership elsewhere.

Using a benchmark-free classification of index-based investors and a global panel of portfolio holdings, we document substantial growth in index ownership across global equity markets, with the largest increase and highest levels in the United States. Value-weighted index ownership increased from 5.1% to 26.2% in the U.S., from 1.8% to 13.9% in Europe, and from 0.8% to 7.5% in Asia between 2004 and 2023. The composition of this growth differs across regions. In the U.S., the increase is driven predominantly by explicit indexing, while in Europe and Asia explicit and closet index ownership have both grown substantially. Explicit indexing surpassed closet indexing in Europe and Asia only recently.

The U.S. also differs from other markets in the cross-section of index ownership. Index ownership extends much further down the market-cap distribution and is highest among mid-cap firms, whereas in Europe and Asia it increases monotonically with firm size and remains concentrated among larger firms. At the firm level, market capitalization and free float are the strongest predictors of index ownership, while trading volume, dividend status, and profitability add relatively little explanatory power. Differences in firm size and free float account for part, but not all, of the substantially higher index ownership observed in the U.S.

These differences matter for interpreting the predominantly U.S.-based literature on

index ownership and passive investing. Our results show that the U.S. is unusual both in the level of index ownership and in which firms are held by index investors, so effects estimated in the U.S. need not generalize to other markets. The global panel developed in this paper makes it possible to examine whether the effects of index ownership on corporate governance, liquidity, price formation, financing, and other firm outcomes vary across countries and with the level and composition of index ownership.

References

Aghion, Philippe, John Van Reenen, and Luigi Zingales, "Innovation and institutional ownership," *American Economic Review*, 2013, *103* (1), 277–304.

Antón, Miguel, Florian Ederer, Mireia Giné, and Guillermo Ramirez-Chiang, "Common ownership around the world," Technical Report, National Bureau of Economic Research 2025.

Appel, Ian R, Todd A Gormley, and Donald B Keim, "Passive investors, not passive owners," *Journal of Financial Economics*, 2016, *121* (1), 111–141.

Bjerksund, Petter, Trond Døskeland, André Wattø Sjuve, and Andreas Ørpetveit, "Forced to be active: Evidence from a regulation intervention," *Management Science*, 2026, *72* (5), 4341–4358.

Boone, Audra L. and Joshua T. White, "The effect of institutional ownership on firm transparency and information production," *Journal of Financial Economics*, 2015, *117* (3), 508–533.

Busse, Jeffrey A. and Qing Tong, "Mutual fund industry selection and persistence," *The Review of Asset Pricing Studies*, 2012, *2* (2), 245–274.

Chinco, Alex and Marco Sammon, "The passive ownership share is double what you think it is," *Journal of Financial Economics*, 2024, *157*, 103860.

Coles, Jeffrey L., Davidson Heath, and Matthew C. Ringgenberg, "On index investing," *Journal of Financial Economics*, 2022, *145* (3), 665–683.

Cremers, KJ Martijn and Antti Petajisto, "How active is your fund manager? A new measure that predicts performance," *The Review of Financial Studies*, 2009, *22* (9), 3329–3365.

Cremers, Martijn, Miguel A Ferreira, Pedro Matos, and Laura Starks, "Indexing and active fund management: International evidence," *Journal of Financial Economics*, 2016, *120* (3), 539–560.

European Fund and Asset Management Association, "EFAMA Fact Book 2024: Trends in European Investment Funds," 2024.

Ferreira, Miguel A. and Pedro Matos, "The colors of investors' money: The role of institutional investors around the world," *Journal of Financial Economics*, 2008, *88* (3), 499–533.

Frazzini, Andrea, Jacques Friedman, and Lukasz Pomorski, "Deactivating active share," *Financial Analysts Journal*, 2016, *72* (2), 14–21.

Haddad, Valentin, Paul Huebner, and Erik Loualiche, "How competitive is the stock market? Theory, evidence from portfolios, and implications for the rise of passive investing," *American Economic Review*, 2025, *115* (3), 975–1018.

Investment Company Institute, "Investment company fact book - A review of trends and activities in the investment company industry," 2024.

Kontz, Christian and Sebastian Hanson, "The real cost of benchmarking," *Working Paper*, 2025.

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

Rowley, James J. Jr., Stephen Lawrence, and Ollie Ryder-Green, "Setting the record straight: The truths about index fund investing," Technical Report, Vanguard August 2025.

Schmidt, Cornelius and Rüdiger Fahlenbrach, "Do exogenous changes in passive institutional ownership affect corporate governance and firm value?," *Journal of Financial Economics*, 2017, *124* (2), 285–306.

Appendix A. Variable Definitions

Variable	Description
$I_{it}^{\text{Index Fund, Morningstar}}$	Indicator equal to 1 if Morningstar classifies fund i as an *index* fund in quarter t.
Turnover_{it}	Gross trading relative to portfolio value: $$\text{Turnover}_{it} \;=\; \frac{\text{Buys}_{it} + \text{Sells}_{it}}{\text{AUM}_{it}},$$ where holdings are reported at the end of each quarter. Buys_{it} and Sells_{it} are defined as all increases and decreases in the fund's holdings between the end of quarter $t-1$ and the end of quarter t (including newly added and fully liquidated positions). AUM_{it} denotes the fund's *end-of-quarter* portfolio value at t.
$\text{Portfolio Holdings}_{it}$	Number of distinct securities held by fund i in quarter t: $$\text{Portfolio Holdings}_{it} \;=\; \big\lvert \mathcal{P}_{i,t} \big\rvert, \quad \mathcal{P}_{i,t} \equiv \{j : \text{SharesOwned}_{ijt} > 0\}.$$
Variance_{it}	Within–portfolio dispersion of free–float–adjusted ownership shares, computed after symmetric 10% truncation at the security cross–section. For $j \in \mathcal{P}_{i,t}$ define $$w_{ijt} \;=\; \frac{\text{SharesOwned}_{ijt}}{\text{FreeFloatShares}_{jt}}, \qquad \bar{w}_{it} \;=\; \text{mean}_{j \in \mathcal{P}_{i,t}}\big(w_{ijt}\big),$$ normalize $\tilde{w}_{ijt} = w_{ijt}/\bar{w}_{it}$, truncate $\tilde{w}_{ijt}$ at the 10th/90th percentiles across j to obtain $\tilde{w}_{ijt}^{W}$, then $$\text{Variance}_{it} \;=\; \frac{\text{sd}_{j \in \mathcal{P}_{i,t}}\big(\tilde{w}_{ijt}^{W}\big)}{\text{mean}_{j \in \mathcal{P}_{i,t}}\big(\tilde{w}_{ijt}^{W}\big)}.$$
$\text{Position Changes}_{it}$	Cross–sectional dispersion of proportional changes in w_{ijt} across the union of current and prior holdings $\mathcal{Q}_{i,t} = \mathcal{P}_{i,t} \cup \mathcal{P}_{i,t-1}$. Let $$\Delta w_{ijt} \;=\; w_{ijt} - w_{ij,t-1}, \qquad r_{ijt} \;=\; \frac{\Delta w_{ijt}}{w_{ij,t-1}}, \quad j \in \mathcal{Q}_{i,t} \text{ with } w_{ij,t-1} > 0.$$ Normalize by the cross–sectional mean magnitude of nonzero–prior positions and take a coefficient of variation: $$\text{Position Changes}_{it} \;=\; \frac{\text{sd}_{\,j:\, w_{ij,t-1}>0}\big(r_{ijt}\big)}{\text{mean}_{\,j:\, w_{ij,t-1}>0}\big(\lvert r_{ijt} \rvert\big)}.$$ If $\text{mean}_{\,j:\, w_{ij,t-1}>0}\big(\lvert r_{ijt} \rvert\big) = 0$, set $\text{Position Changes}_{it} = 0$. Equals 0 if all nonzero prior positions are scaled proportionally between $t-1$ and t.
AUM_{it}	Portfolio market value of fund i in quarter t; used to scale Turnover_{it} and to compute position weights.
$\text{Index BD or Name}_{i}$	Time-invariant indicator equal to 1 if the fund's business description indicates index tracking or replication; if the business description is unavailable, equals 1 when the fund name contains any of: *Index, CAC, DAX, OMX, FTSE, S&P, S&P 500, SP500, MSCI, Nasdaq, Russell, Total Market, QQQ, SPY, IVV, VOO, VTI, IWM, Set 50, Jumbo 50, Jumbo 25, Select Sector SPDR, Euro Stoxx, Topix, Nikkei, JASDAQ, IBEX, TSX, ASX, Bovespa, KOSPI, Hang Seng, HSI, Lyxor Index, Dow, iShares Core, Vanguard Index, Vanguard 500, Vanguard Total Stock, Xtrackers, Schwab 1000, Fidelity 500*.
FreeFloat_{jt}	Firm–level free float for security j in quarter t: the proportion of shares available for public trading, excluding shares held by strategic owners. Used to construct w_{ijt} in Variance_{it} and $\text{Position Changes}_{it}$.

Table A1: Variable definitions and construction. This table provides definitions for key variables used in the paper. Indices are i (fund), j (firm/security), and t (calendar quarter).

Appendix B. Examples of Fund Business Descriptions

Below, we present three examples of the business descriptions used in our classification procedure. The first description refers to indexes as performance benchmarks only, whereas the last two describe index-tracking strategies and replication techniques. The descriptions are the latest available (as of 2026).

Example 1: Pembroke Canadian Growth Fund

> Pembroke Canadian Growth Fund is an open-end equity mutual fund launched by GBC Asset Management Inc. It is managed by Pembroke Management Ltd. The fund invests in public equity markets of Canada. It invests in stocks of companies operating across diversified sectors. The fund invests in growth and value stocks of companies across diversified market capitalization. It employs fundamental analysis with bottom-up stock picking approach to create its portfolio. The fund seeks to benchmark the performance of its portfolio against the S&P/TSX Total Return Index and the S&P/TSX Small Cap Index. Pembroke Canadian Growth Fund was formed on September 8, 1988 and is domiciled in Canada.

Example 2: iShares MSCI Switzerland ETF

> iShares, Inc. - iShares MSCI Switzerland ETF is an exchange traded fund launched by BlackRock, Inc. The fund is managed by BlackRock Fund Advisors. It invests in public equity markets of Switzerland. The fund invests in stocks of companies operating across diversified sectors. The fund invests in growth and value stocks of companies across diversified market capitalization. The fund seeks to track the performance of the MSCI Switzerland 25/50 Index, by using representative sampling technique. iShares, Inc. - iShares MSCI Switzerland ETF was formed on March 12, 1996 and is domiciled in the United States.

Example 3: Invesco Bloomberg MVP Multi-factor ETF

Invesco Exchange-Traded Fund Trust - Invesco Bloomberg MVP Multi-factor ETF is an exchange traded fund launched and managed by Invesco Capital Management LLC. The fund invests in public equity markets of the United States. It invests in stocks of companies operating across diversified sectors. The fund invests in momentum, value and volatile stocks of large-cap companies. It seeks to track the performance of the Bloomberg MVP Index, by using full replication technique. Invesco Exchange-Traded Fund Trust - Invesco Bloomberg MVP Multi-factor ETF was formed on May 1, 2003 and is domiciled in the United States.

About ECGI

ECGI is an international non-profit membership organisation that provides a platform for debate and dialogue among academics, policymakers, and business leaders, with a focus on major corporate governance, ESG, and stewardship issues.

The views expressed in this working paper are those of the authors, not those of the ECGI or its members.

www.ecgi.global

ECGI Working Paper Series in Finance

https://www.ecgi.global/publications/working-papers