Research-Thesis: A comparative performance analysis of quantitative and discretionary approaches in investment fund management.

Auteur : Devresse, Pierre
Promoteur(s) : Bodson, Laurent
Faculté : HEC-Ecole de gestion de l'Université de Liège
Diplôme : Master en sciences de gestion, à finalité spécialisée en Banking and Asset Management
Année académique : 2025-2026
URI/URL : http://hdl.handle.net/2268.2/25824

HEC LIÈGE
Management School - Liège Université

A COMPARATIVE PERFORMANCE ANALYSIS OF QUANTITATIVE AND DISCRETIONARY APPROACHES IN INVESTMENT FUND MANAGEMENT

Jury:
Supervisor:
Laurent BODSON
Reader(s):
Thi Thuy Van NGUYEN

Master thesis presented by
Pierre DEVRESSE
To obtain the degree of
MASTER IN MANAGEMENT
with a specialization in
Banking and Asset Management
Academic year 2025/2026

Ecole de Gestion de l'Université de Liège

Acknowledgments

My first thanks go to my supervisor, Professor Laurent Bodson, who guided this thesis from its earliest and roughest outline through to its final form. He pushed me to justify each methodological choices instead of settling for the first answer that came to mind. I know for a fact that he is really busy, and I am grateful he accepted to assist me in the completion of this milestone.

My thanks extend to Thi Thuy Van Nguyen, who agreed to serve as reader for this work.

To my family, my girlfriend and closest friends, thank you for your patience over what became a long journey. You were always there to get me back on my feet when I was in doubt or lost. Always trying to help me to the best of your abilities with this complicated subject. I could not have reached the end of this without you.

Finally, I want to thank the professors and staff of HEC Liège, who gave me the grounding to take on a project of this scope and made a long degree a rewarding experience. I leave grateful for those years.

Table of contents

1. Glossary

1. AI — Artificial intelligence
2. AUM — Assets under management
3. CAPM — Capital Asset Pricing Model
4. CMA — Conservative minus aggressive (investment factor)
5. ESG — Environmental, social and governance
6. FF6 — Fama and French 5 factors model + momentum
7. HML — High minus low (value factor)
8. LSEG — London Stock Exchange Group (provider of LSEG Workspace, formerly Eikon)
9. MKT — Market factor (excess return on the market portfolio)
10. ML — Machine learning
11. MOM — Momentum factor (past twelve-month winners minus losers)
12. NAV — Net asset value
13. OLS — Ordinary least squares
14. PIS — Principal investment strategy (section of a fund prospectus where the investment decision process is describe)
15. RI — Return index (total-return series)
16. RMW — Robust minus weak (profitability factor)
17. SAFE — Sustainability, Accuracy, Fairness and Explainability (Giudici & Raffinetti, 2023)
18. SMB — Small minus big (size factor)
19. TER — Total expense ratio
20. TNA — Total net assets
21. US — United States
22. USD — US dollar
23. VIF — Variance inflation factor

2. Introduction

2.1. Motivation

Over the past two decades, the asset management industry has experienced a profound transformation. Two major forces have reshaped the way investment decisions are made: the rapid adoption of computerised model in finance and the unprecedented growth of available data. These developments are not just refining existing practices, they are also questioning the traditional boundary between human judgement and algorithmic decision-making, which raises questions of considerable importance for both academics and practitioners.

The first of these forces is the speed at which computerised models have spread across the industry. From the early adoption of systematic factor strategies in the 1990s to today's machine learning-based trading systems, quantitative methods have progressively taken over parts of the investment process that used to be the exclusive domain of human managers. In the United States, this evolution is particularly striking. According to The Economist (2019), quantitative funds now represent approximately 35% of stock market ownership, around 60% of institutional equity assets under management and 60% of trading volume. It is worth noting that those numbers are inherently dependent on the classification of funds as quantitative and the scope of market participants considered. Miguel and Chen (2025) document that the number of US-domiciled funds classified as quantitative by Lipper more than quadrupled between 2000 and 2019, increasing from 35 to 163 funds in their sample. A similar dynamic can be observed in China, where the assets under management of quantitative funds grew massively from 2014 onwards (Hu et al., 2024). Systematic investing has therefore become a structural strategy of global equity markets, and no longer a marginal phenomenon.

The second force is the explosion of big data which is combined with the rise in computational power that makes it processable. The volume and variety of financial data available to fund managers has grown at an extraordinary rate. Today, this data includes not only traditional financial statements and metrics but also satellite imagery, social media sentiment, consumer transaction records and textual disclosures (Zhang et al., 2023). At the same time, the sharp decrease in computing costs has made it possible to deploy non-linear machine learning models, such as random forests, gradient boosting and deep neural networks, on a large scale. For instance, DeMiguel et al. (2023) show that these models significantly outperform classical linear regressions when predicting future mutual fund alpha based on their characteristics. Together, the abundance of data and computational power have given

quantitative approaches analytical possibilities that would have been unimaginable when Markowitz (1952) first formalised modern portfolio theory.

However, the superiority of systematic methods is far from obvious, and this is precisely where the debate becomes interesting. On the theoretical side, quantitative approaches present real advantages. By following predefined rules, they can largely avoid the behavioural biases that affect even professional investors, such as overconfidence, the disposition effect[1] or loss aversion (Dyer et al., 2025). They also offer a level of consistency, objectivity and scalability that human managers cannot reasonably match across thousands of securities at the same time. On the other hand, rule-based algorithms are built on historical data and predefined patterns. When markets experience structural change or when the informational environment shifts, these patterns may no longer hold. Dyer et al. (2025) provide a particularly clear illustration of this limitation: following the implementation of major new accounting standards, quantitative funds underperformed their discretionary counterparts by up to 30 basis points per month, as their models failed to reinterpret the new information in the way a human analyst could.

This creates an important and unresolved trade-off between the two approaches. Quantitative managers excel at processing large volumes of hard, machine-readable data in a consistent and scalable manner (Miguel & Chen, 2025). Discretionary managers, on the other hand, retain a comparative advantage in the interpretation of soft information, that is, qualitative elements such as the assessment of management quality, industry dynamics or strategic context (Birru et al., 2024). They also tend to react more flexibly to new situations for which no historical precedent exists (Dyer et al., 2025; Miguel & Chen, 2025). A further complication concerns the transparency of the two approaches. As quantitative models become more complex, their decision-making process increasingly become what is caled a black box, which raises legitimate concerns regarding trust, accountability and explainability for both investors and regulators (Habbal et al., 2024). In response, frameworks such as the SAFE model proposed by Giudici and Raffinetti (2023) have been developed to evaluate AI systems along four key dimensions: Sustainability, Accuracy, Fairness and Explainability. Discretionary management, despite its subjective dimension, at least provides an explainable human rationale.

Despite the growing empirical interest in this debate, the existing literature remains fragmented. Studies tend to focus on specific markets, narrow time windows or particular technologies. For instance, Chen and Ren (2022) study AI-powered funds in the United States, Zhang et al. (2023) examine big data funds in China, while Chincarini (2014) focuses on hedge funds. The results obtained

[1] The disposition effect is defined as the tendency to hold onto losers and sell winners. See Cici (2012) for more information on this effect for mutual funds.

across these studies are often contradictory and difficult to reconcile. A comprehensive, long-horizon comparison based on a broad sample and a robust multi-factor framework is still largely missing from the literature. This thesis aims to contribute to filling this gap.

2.2. Research Objective

The increasing adoption of quantitative and data-driven techniques in asset management has generated substantial academic interest, particularly with respect to their implications for mutual fund performance. This master thesis aims to contribute to this literature by examining how the two different decision-making approaches relate to performance outcomes.

The primary objective of this research is to compare the performance of quantitative funds against their discretionary counterparts. While both types of funds can involve human implication and the use of computational tools, they differ fundamentally in the degree to which portfolio decisions are governed by managerial discretion versus computerised models. These differences may lead to distinct performance characteristics, which motivates the first and central research question of this study:

RQ1: Do quantitative funds outperform discretionary funds on a risk-adjusted basis?

This question is addressed over a long and economically varied sample period, spanning from 2000 to 2024 and including major episodes of market stress such as the dot-com collapse, the 2008 financial crisis and the COVID-19 shock. The advantage of treating such a long window is that it allows us to observe both stable and turbulent environments. This is particularly relevant because the existing literature suggests that the relative performance of the two approaches may not be constant over time. Quantitative funds tend to perform well in data-rich and stable environments, while discretionary managers may recover ground during periods of structural change or informational disruption (Dyer et al., 2025; Miguel & Chen, 2025).

Beyond the question of whether one approach outperforms the other, a second question explores their underlying drivers. In particular, this study examines whether differences in fees, factor exposures and fund characteristics help explain the relative outcomes observed across the two management styles. Which leads to the second research question:

RQ2: In which specific dimensions does each management style hold a comparative advantage over the other?

Together, these two research questions aim to go beyond a simple verdict on which approach is "better" and to produce instead a more nuanced picture of when and why systematic discipline or human judgement creates value. By applying a rigorous six-factor performance evaluation framework to a broad sample of equity mutual funds, this study seeks to provide fresh empirical evidence on a central and enduring debate in modern asset management.

3. Literature Review

3.1. The Discretionary Investment Approach

3.1.1. Definition of the Discretionary Approach

What sets discretionary investment management apart from its quantitative counterpart is, first and foremost, the nature of the information it processes and the way investment decisions are reached. Rather than following pre-defined algorithmic rules, a discretionary manager constructs an investment thesis by synthesizing a wide and often heterogeneous variety of inputs: financial statements, management assessments, competitive landscape analyses, macroeconomic outlooks, and — perhaps most distinctively — soft information that is hard to quantify. As Birru et al. (2024, p. 101688) note, this includes a comprehensive study of "companies' business models, industry dynamics, and strategic choices, as well as quality of management and soft information." Private conversations with corporate executives, on-site company visits, and the accumulated knowledge of years of sector specialization are all part of the discretionary manager's toolkit. These are dimensions of the investment process that are, by their very nature, difficult to replicate systematically.

This approach finds part of its justification in the theoretical foundation of Grossman and Stiglitz (1980), who argued that markets cannot be fully informationally efficient if there is no financial reward to acquiring and acting on private information. In their framework, some degree of informed trading is a necessary condition for prices to reflect fundamental values at all. Discretionary managers, particularly those with deep sector expertise, can be understood as these informed agents — participants whose specialized knowledge contributes to keep prices approximately efficient while allowing them to extract a temporary informational rent[2].

[2] An informational rent is the economic profit that an informed agent earns because they possess information that is not yet reflected in market prices. In the Grossman-Stiglitz framework, this rent compensates investors for the costs they incur in acquiring and processing private information. If markets were perfectly efficient and all information were already in prices, such compensation would not exist, and rational investors would stop gathering information.

3.1.2. The Role of Managerial Skill and Industry Specialization

The question of whether the returns generated by discretionary managers genuinely reflect skill — as opposed to risk-taking, luck, or style attributes — has attracted a vast body of academic investigation. Within this literature, a growing body of evidence suggests that when skill does exist, it tends to be concentrated among managers with deep, sector-specific expertise rather than generalist market knowledge

Jordan et al. (2022) offer one of the most recent demonstrations of this mechanism. Studying the portfolio construction choices of active mutual fund managers, they show that those with specialized industry knowledge exhibit a pronounced preference for pure-play firms[3] over their diversified conglomerate peers. The key empirical contribution of their work is that a fund's degree of industry concentration, which they defined as a "pureplayness" measure, is a statistically and economically significant predictor of before-fee alpha, industry selectivity, and industry timing ability. Their results suggest that managers who concentrate their portfolios around firms they truly understand are able to translate that understanding into outperformance — not through broad market calls, but through precise, specialized security selection.

This finding connects with a broader theoretical argument: that discretionary value-added, likely arises from the processing of information that is either private, soft, or too context-dependent to be captured by a statistical model. Grinblatt and Titman (1994) provided early evidence of positive gross abnormal returns for a subset of equity mutual funds, interpreting their findings as consistent with genuine information advantages for at least some managers. More recently, Coleman (2023) frames the same intuition through the lens of the structure-conduct-performance paradigm[4], arguing that fund managers who operate in niches with genuine informational barriers — where their private network and sector knowledge truly differentiate them — may sustain real outperformance before fees.

[3] A pure-play firm is a company that derives essentially all of its revenue from a single industry or line of business — as opposed to a diversified conglomerate operating across multiple sectors. For fund managers, pure-play firms are attractive precisely because their financial performance maps directly onto industry-specific dynamics, making them easier to analyse with deep sector expertise

[4] The *structure-conduct-performance* (SCP) paradigm is a framework from industrial organization economics (Bain, 1959) which posits that an industry's structural characteristics (concentration, barriers to entry) shape the behaviour of firms within it, which in turn determines their performance outcomes. Coleman (2023) applies this logic to the mutual fund industry, arguing that its oligopolistic structure and AUM-based fee models incentivize conduct — such as herding and window dressing — that systematically undermines fund performance.

3.1.3. Structural and Behavioural Limitations

Despite the theoretical appeal of human expertise as a source of alpha, a substantial body of research documents a persistent gap between the promise of active discretionary management and its delivered outcomes. This gap has been attributed to two broad sets of forces: behavioural imperfections at the level of individual managers, and structural conflicts of interest embedded in the institutional design of the fund management industry.

On the behavioural side, Verma et al.(2024) provide rich qualitative evidence, drawn from direct interviews with active fund managers, that emotional biases remain present even among experienced professionals. Their respondents acknowledge being affected by overconfidence, loss aversion, and cognitive anchoring[5] — the biases that the behavioural finance field has long argued distort human judgment under uncertainty. These confessions are notable not because they are surprising on their own, but because they come from practitioners who are presumably aware of these biases and actively trying to manage them. If even experienced managers struggle to eliminate emotional interference, the question arises as to whether human judgment is systematically disadvantaged relative to computerised processes in environments that demand consistent, dispassionate decision-making.

On the structural side, the problems are arguably more fundamental. Mahoney (2004) identifies the mismatch between managerial incentives and investor interests as a defining feature of the mutual fund industry. The dominant fee model — in which managers are compensated primarily as a percentage of assets under management rather than as a share of net returns — creates incentives to prioritize fund growth over investment performance. This problem is not only theoretical: Mahoney documents how it was a root cause of several of the conflicts of interest that erupted in U.S. mutual fund scandals in the early 2000s, where managers' AUM-linked compensation motivated practices such as market timing and late trading[6] that systematically disadvantaged long-term investors.

Yin (2016) extends this analysis to the hedge fund context, where incentive fee structures are generally considered better aligned with investor interests. Even here, he shows empirically that managerial compensation is maximized at a fund size considerably larger than the size that would optimize risk-

[5] Cognitive anchoring is a well-documented behavioural bias. It describes the tendency of individuals to rely too heavily on an initial piece of information (the "anchor") when making subsequent judgments, even when the anchor is arbitrary or outdated. In an investment context, a fund manager anchored to a prior valuation target may be slow to update their view considering new information, leading to systematic mispricing in their portfolio.
[6] Market timing, in this context, refers to the practice of allowing favoured investors to buy or sell fund shares based on knowledge of anticipated portfolio changes or market movements, at the expense of other shareholders. Late trading refers to the illegal practice of placing orders to buy or redeem fund shares after the market close at 4:00 p.m. while still receiving the same day's net asset value — effectively allowing certain investors to trade on information not yet reflected in the fund's price

adjusted performance. The implication is important: even in environments designed to better align incentives, structural forces push against pure performance-maximization.

Coleman (2023) argues that these structural problems are systemic rather than incidental. The industry's oligopolistic structure, combined with information asymmetries that make it difficult for unsophisticated investors to evaluate true fund quality, creates an environment where managers face only weak competitive pressure to deliver genuine outperformance. Herding behaviour — where managers mimic each other's portfolio choices to reduce career risk — emerges naturally from this structure, as the professional cost of underperforming peers tends to be more damaging than the cost of underperforming the market when everyone else does too.

The interaction between fund size and performance deserves particular attention. Pástor, Stambaugh and Taylor (2015) provide rigorous empirical evidence that returns to scale in active management are strongly negative at the industry level. As the active fund industry grows, price impact rises, available investment opportunities shrink, and transaction costs per unit of alpha deteriorate. Their estimates suggest that the industry as a whole operates beyond its efficient scale, implying persistent overcrowding. This dynamic is self-reinforcing: because capital flows to past winners, successful funds grow beyond their optimal size, and their performance tends to revert. Berk and Green (2004) formalize this mechanism in a rational equilibrium model, showing that the absence of persistent gross-of-cost outperformance is fully consistent with the existence of skilled managers, once the equilibrating role of fund flows is accounted for. The key insight is that skilled managers capture the benefits from their skill through fees rather than through delivered net-of-fee returns.

3.1.4. The Empirical evidence on Active Discretionary Performance

The accumulated empirical evidence on net-of-fee performance of discretionary mutual funds is overall discouraging. The study by Fama and French (2010), which applies a bootstrap methodology[7] to isolate luck from skill across the full cross-section of U.S. equity mutual fund returns, finds that the distribution of alphas is almost entirely consistent with a world in which no manager has any skill at all. Only the very right tail of the performance distribution — roughly corresponding to the top 3% of

[7] In this context, the bootstrap methodology is a simulation-based statistical technique used to construct a distribution of fund performance under the assumption of zero managerial skill. Fama and French (2010) generate thousands of simulated fund return histories by resampling from the actual data while imposing the constraint that no fund has any genuine alpha. The observed distribution of real fund alphas is then compared against this simulated benchmark: funds whose actual performance exceeds what the simulation produces at the 97th percentile are candidates for real skill, while those within the simulated range could be explained by chance alone

funds — have returns that cannot be explained by random variation alone. Most actively managed, discretionary funds deliver negative risk-adjusted returns after fees, confirming what Jensen (1968) had already documented five decades earlier using a simpler CAPM framework.

This conclusion is, of course, an average. It does not exclude the existence of genuinely skilled individual managers — and, as the evidence on sector specialization discussed above suggests, such managers do appear to exist. What it does establish is that identifying them is a really difficult task.

3.2. The Quantitative Investment Approach

3.2.1. Defining the Quantitative Approach and Its Foundations

A fund is generally described as quantitative when the choice of which securities to hold, and in what proportion, is delegated to computer-based models rather than decided by a manager (Dyer et al., 2025; Miguel & Chen, 2021, 2025). The distinction resides in decision authority rather than tooling: discretionary managers also rely on screens, spreadsheets and statistical software, but they remain free to ignore what those tools suggest. In a quantitative fund, once the model has been specified and its inputs chosen, the portfolio produce mechanically its output. Human involvement does not disappear — it moves upstream, into the design of the model, the selection of signals and the supervision of execution — but it is deliberately removed from the individual buy-and-sell decision.[8] Underlying this design is a fundamental trade-off: computing power gives systematic approaches a greater capacity for processing information at scale, while their reliance on fixed rules limits their ability to adapt to conditions that fall outside historical patterns. This master's thesis adopts the broad definition used in the recent literature and treats as quantitative any fund whose disclosed investment process is driven primarily by systematic and computer-based models, without distinction on the signals they exploit.

The mathematical foundations of this approach predate the computing power that made it practical. Markowitz (1952) was the first to give portfolio construction a formal language, showing that the risk of a portfolio depends not on the riskiness of its individual holdings but on the covariances between them, so that diversification becomes a quantifiable property rather than a vague principle of

[8] In practice, the classification of a fund as quantitative or discretionary is less straightforward than the definition implies, since it resides primarily on prospectus language which is often imprecise and the boundary between the two styles has blurred as discretionary managers adopt quantitative tools. Abis (2020) develops a text-based classification method to address this difficulty. She also explicitly articulates the capacity versus flexibility trade-off underpinning the distinction: computing power gives quantitative funds superior information processing capacity, while their reliance on fixed rules limits their adaptability to structural shifts in market conditions

prudence. The object of interest thereby shifted from the single security to the portfolio as a whole. The framework was appealing but fragile in application: Black and Litterman (1992) observed that feeding estimated expected returns directly into a mean-variance optimiser typically produced portfolios with large, unstable and economically implausible positions, because the optimiser reacts sharply to small estimation errors. Their solution was to start from the market-capitalisation-weighted portfolio as a neutral equilibrium and to deviate from it only to the extent that the manager holds an explicit, quantified view. That compromise — a model disciplined by a market anchor yet still open to judgment expressed in a structured form — captures well the spirit in which most quantitative strategies are actually applied in practice.

3.2.2. Factor Investing and the Proliferation of Signals

If modern portfolio theory supplied the language, factor investing supplied the dominant working method. Rather than attempting to identify individual mispriced stocks, a factor investor ranks the whole investable universe on a measurable characteristic and holds a portfolio designed to earn the premium historically associated with it. Asness et al. (2013) make a particularly influential case for two of them, showing that the value and momentum premia appear consistently across eight markets and asset classes, from individual equities in several countries to currencies, bonds and commodities. The strength of that result lies in its large scope: a premium visible in so many distinct settings is unlikely to be the statistical accident of a single dataset, and more plausibly reflects a genuine economic phenomenon.

The same empirical process, however, proved to be dangerous. Harvey et al. (2016) surveyed the academic literature and counted more than 300 published factors — a figure they interpreted as implausible and largely the product of data mining, since with enough candidate variables some will appear significant by chance alone.[9] Their recommendation was to raise the bar, treating a t-statistic close to 3.0, rather than the conventional 2.0, as the minimum for taking a new factor seriously. This concern is backed by a related finding from Mateus et al. (2019), whose review of the factor-model literature concludes that no single proposed factor has proven able to explain all anomalies or price all assets. Their review further shows that the choice of benchmark materially affects the performance attributed to a fund: a model that omits a premium, a fund is mechanically harvesting will misattribute that risk exposure as alpha.

[9] A factor identified this way is typically the product of backtesting, evaluating a rule based on historical data, and is exposed to overfitting: the rule may be tuned so closely to the past that it captures noise rather than a stable relationship and fails out of sample (Harvey et al., 2016)

3.2.3. The Machine-Learning and A.I. revolution

The methods described above share a common limitation: the researcher must specify in advance the functional form linking signals to returns, almost always a linear one. The most consequential recent development in quantitative investing has been the adoption of techniques that lift this constraint. Gu et al. (2020) provide one of the reference study, comparing a broad set of machine-learning methods on the task of predicting stock returns. They find that the more flexible methods — in particular regression trees[10] and neural networks[11] — roughly double the out-of-sample performance of established linear strategies, and that the gain comes from capturing interactions and non-linearities that a linear specification cannot represent. Comfortingly, the methods broadly agree on which signals matter most, namely variants of momentum, liquidity and volatility, which suggests they are extracting economic structure rather than noise. DeMiguel et al. (2023) show that these non-linearities extend to fund selection as well: machine-learning (ML) models exploiting fund characteristics identify future outperformers with sufficient precision to generate annual alphas of 2.4% net of all costs.

Machine learning has since spread well beyond return forecasting. Creamer and Freund (2010) built a fully automated trading systems, combining a boosting algorithm[12] with decision trees to generate signals from technical indicators. Zhang and Huang (2021) apply recurrent neural networks to the rather different problem of options hedging, with results that improve on conventional delta-hedging. More broadly, these approaches appear to reduce the susceptibility to cognitive biases that affect even experienced discretionary managers: Chen and Ren (2022) find that AI-powered funds outperformance is partly attributable to reduced behavioural biases, a structural advantage also documented in the hedge fund context by Chincarini (2014).

A similar part of the literature has explored whether gains from alternative data sources — satellite imagery, social-media activity, transaction records — match those from better algorithms. Here the

[10] A regression tree is a decision-making algorithm that works like a flowchart: it repeatedly splits the data into two groups based on whichever variable best separates high-return stocks from low-return ones, until it arrives at a prediction. The advantage over a standard regression is that it detects complex patterns and combinations of signals automatically, without the researcher having to anticipate them. In practice, hundreds of such trees are built simultaneously and their predictions averaged, which produces a more reliable result than any single tree (Gu et al., 2020)

[11] A neural network is inspired by the structure of the human brain: it consists of layers of interconnected nodes that each transform their inputs and pass the result forward. By stacking many such layers, the network can learn extremely complex patterns in the data. In a financial context, the inputs are assets characteristics and the output is a return prediction; the network figures out on its own which combinations of signals matter, without being told what to look for (Gu et al., 2020).

[12] A boosting algorithm builds a prediction by combining many simple, individually weak models in sequence. Each new model focuses specifically on the cases where the previous ones made the largest errors, so that the ensemble progressively corrects its own mistakes. The result is a powerful predictor that emerges from the accumulated corrections of many small improvements rather than from a single complex model (Creamer & Freund, 2010).

evidence is more discouraging: Zhang et al. (2023) find that big-data Chinese mutual funds do not outperform their human-managed peers, and that the big-data factor does not meaningfully improve stock selection. The difficulty of replicating model-level ML performance in fund management is a recurring theme and reinforces the interpretability concern. The most powerful models are difficult to explain in human digestible terms — a central concern in a regulated industry where decisions must be justified to clients and supervisors[13]. Frameworks designed to address this tension, such as the SAFE approach of Giudici and Raffinetti (2023), have begun to emerge, though the trade-off between predictive accuracy and accountability is unlikely to be resolved in the near term.

3.2.4. Market Efficiency and the Hybrid Frontier

Beyond its effects at the fund level, systematic investing also reshapes the market in which it operates. Because quantitative strategies trade on the patterns that the literature documents, their activity tends to push prices toward fundamental value. Birru et al. (2024) show that quantitative sell-side analysts help other investors recognise and act on anomaly signals, thereby contributing to more efficient pricing. Studying the rapid expansion of quantitative funds in China, Hu et al. (2024) reach a convergent conclusion: as systematic capital grows, the returns to well-known anomalies shrink and the market becomes more efficient. The result carries an irony that matters for the remainder of this master's thesis: the more successfully quantitative investors exploit a signal, the more they reduce the very premium that justified it — which helps explain both the relentless search for new signals documented in Section 3.2.2 and the difficulty of sustaining outperformance over time. Part of the same dynamic is the overcrowding that emerges when many quantitative funds draw on the same published signals and common data vendors, generating correlated positions that can revert sharply under stress.[14]

It would nonetheless be a mistake to read this evolution as a straightforward replacement of human judgment by machines. One of the most recent evidence points instead toward complementarity. Cao et al. (2024) compare an AI analyst with human equity analysts and find a clear division of labour: the machine is superior when information is transparent but voluminous, whereas humans retain an advantage when the relevant knowledge is intangible or context-dependent, as in situations of

[13] The "black box" expression refers to the difficulty of explaining how a complex model arrives at a prediction. In a linear regression each variable's contribution is an explicit coefficient; a deep neural network offers no comparable account, which is a real constraint where investment decisions must be justified to clients and supervisors (Giudici & Raffinetti, 2023)

[14] Abis (2020) provides direct evidence of this: compared to discretionary funds, quantitative funds hold more stocks but engage in significantly more overcrowded trades. A pattern consistent with their shared reliance on published academic signals and common data sources.

financial distress. Crucially, combining the two — what the authors call "Man + Machine" — outperforms either alone and sharply reduces large forecasting errors. On this evidence, the frontier of quantitative investing is not man against machine but man with machine, a mix in which each side compensates for the other's weakness.

3.3. Performance Evaluation of Investment Funds

Comparing two management styles only becomes meaningful once a credible way of measuring performance has been accepted. A raw return tells us little on its own: a fund may have earned a high return because it was exposed to a rising market, or because it bet heavily on small or value stocks that happened to do well over the period. The question that matters for this master's thesis is therefore not whether quantitative or discretionary funds earned more, but whether either group earned more than its risk exposures would predict. This section reviews how the literature has approached that question and explains why the present study relies on a six-factor Fama–French–Carhart specification estimated with region-matched factors.

3.3.1. Measuring abnormal return: alpha, benchmarks, and the skill–luck problem

The traditional measure of fund performance is Jensen's alpha — the intercept of a regression of a fund's excess return on the excess return of a benchmark portfolio (Jensen, 1968; Kuosmanen, 2007). A positive and significant alpha is interpreted as evidence that the manager added value beyond passive exposure to systematic risk. This definition of skill as a regression intercept has remained the standard in the literature. In practice, alpha is rarely the end of the analysis: a common design first estimates alpha from a time-series factor regression and then, in a second stage, relates it to observable fund characteristics through a cross-sectional regression — the two-step procedure used, among others, by Ferreira et al. (2013), and adopted in the empirical study of this study.

Two difficulties complicate the interpretation of alpha. The first is the hypothesis problem: any statement about abnormal performance is inseparable from the model used to define "normal" performance. Grinblatt and Titman (1994) demonstrate this concretely, showing that performance

inferences are sensitive to the benchmark chosen and that mean-variance inefficient benchmarks[15] can produce misleading conclusions. Kuosmanen (2007) likewise stresses that the alpha estimate remains sensitive to the choice of benchmark portfolio. A negative alpha may therefore reveal a weak manager, or simply a poorly specified model. The second difficulty is statistical: because returns are volatile, a fund can post a run of positive alphas purely by luck. Separating genuine skill from chance is the central concern of Fama and French (2010), whose results, discussed earlier in section 3.1.4, warn against relying too much onto any single positive estimate. From a measurement standpoint, the implication is that the model must be demanding enough to absorb the obvious sources of "normal" return, so that the alpha that remains is a stronger representation of skill.

It is worth recalling why a non-zero alpha should be expected to exist in the first place. Grossman and Stiglitz (1980) argued that perfectly informationally efficient markets are impossible, since no one would gather costly information if prices already reflected it. Their argument leaves room, in equilibrium, for informed managers to be rewarded for the information they produce which is what makes a comparison of discretionary and quantitative skill an empirically meaningful exercise rather than a predestined conclusion.

3.3.2. From the CAPM to the Fama–French–Carhart framework

The single-factor Capital Asset Pricing Model (CAPM) underlying Jensen's (1968) measure is unable to account for several regularities in average returns, which dictated a progressive enrichment of the benchmark. Capocci and Hübner (2004) summarise this standard progression: from the CAPM to the Fama and French (1993) three-factor model, and then to the Carhart (1997) four-factor model. Fama and French (1993) added two factors to the market — SMB (small minus big), which captures the tendency of small-capitalisation stocks to outperform large ones, and HML (high minus low), which captures the outperformance of high book-to-market (value) stocks over growth stocks. Carhart (1997) then isolated one anomaly the three-factor model still could not explain — the momentum effect of Jegadeesh and Titman (1993), by which recent winners keep outperforming recent losers — and added a fourth, momentum factor, reporting that this addition effectively reduced pricing errors relative to the CAPM and the three-factor model. The four-factor model subsequently became the reference benchmark in the mutual fund literature, to the point that Mateus, Mateus, and Todorovic (2019)

[15] A portfolio is mean-variance efficient, in the sense of Markowitz (1952), if it delivers the highest expected return for a given level of risk (variance). In performance evaluation, alpha equals zero if and only if the benchmark is mean-variance efficient (Ferson & Schadt, 1996). When it is not, a manager can appear to outperform purely through exposure to assets the benchmark fails to capture, with no genuine skill involved.

describe the Fama–French and Carhart models as the accepted norm in academic studies of fund performance.

More recently, Fama and French (2015) extended their own framework to five factors by adding RMW (robust minus weak), an operating profitability factor (robust minus weak), and CMA (conservative minus aggressive), an investment factor. The motivation is theoretical as much as empirical: the dividend discount model[16] implies that, holding book-to-market constant, expected returns should be related to profitability and to investment. The five-factor model improves the description of average returns, although Fama and French (2015) acknowledge that it still fails to explain the low returns of small, unprofitable firms that nonetheless invest aggressively, and that HML becomes largely redundant once RMW and CMA are included.

3.3.3. The six-factor model adopted in this master's thesis

One feature of the five-factor model is directly relevant to the design of this study: Fama and French (2015) deliberately left momentum out. This is a notable omission for the evaluation of actively managed equity funds. Momentum is among the most persistent anomalies documented in the literature (Asness et al., 2013; Jegadeesh & Titman, 1993), and funds carry measurable exposure to it — Ferreira et al. (2013) report non-zero average momentum loadings for equity funds across the countries in their sample. Leaving momentum out of the benchmark would therefore risk attributing a known, passively available premium to managerial skill. Combining the five-factor model with a momentum factor addresses this concern, and the resulting six-factor (FF6) specification has become a common choice in recent fund-performance research (DeMiguel et al., 2023; Dyer et al., 2025) making it a natural benchmark for the empirical analysis conducted in this master's thesis

The decision to estimate region-matched factors, rather than to apply U.S. factors uniformly, rests on further evidence. Asness et al. (2013) show that value and momentum premia are present not only in U.S. equities but consistently across international stock markets and asset classes, with a strong common structure across countries. Ferreira et al. (2013) go further on the specific question of fund

[16] The dividend discount model establish that a stock's price equals the present value of its expected future dividends. Gordon (1959) provides one of its foundational formulations, showing that under constant growth in dividends the price can be expressed as $P_0 = (1 - b)Y_0 / (k - br)$, where Y_0 is current earnings, b the retention ratio, r the firm's return on investment, and k the market's required rate of return. The growth rate of dividends is therefore br, meaning that higher retained earnings and a higher return on investment both increase the stock's value. Fama and French (2015) exploit a similar framework: rearranging the dividend discount identity reveals that, for a given price-to-book ratio, a firm with higher expected profitability must have higher expected returns — which motivates RMW — and a firm that invests more aggressively must have lower expected returns, all else equal — which motivates CMA

evaluation: in a study of open-end equity funds covering 27 countries, they construct country-specific Carhart factors and find that the four-factor model explains fund returns at least as well outside the United States as within it, reporting an average R^2 of 88% for non-U.S. funds against 85% for U.S. funds. The evidence that a model first developed on U.S. data translate well to other markets is reassuring for this study whose sample is international, and it directly supports matching each fund to the factor set of its own investment region — consistent with the warning of Grinblatt and Titman (1994) that an un-suited benchmark distorts performance inferences. The choice of the six-factor model estimated for each fund in this work, which will be specified in the methodology chapter, is a direct outcome of all these considerations.

3.3.4. Known limitations: unconditional model and the choice of alpha metric

The six-factor model used here is unconditional: it assumes that a fund's factor loadings remain constant throughout the estimation window. Ferson and Schadt (1996) challenged this assumption, showing that fund betas vary predictably over the business cycle. When they do, an unconditional regression can misattribute a mechanical response to changing market conditions as genuine selection skill. Their solution — allowing betas to vary with lagged public information such as dividend yields or short-term interest rates — produces more accurate alpha estimates. This study retains the unconditional approach for simplicity and to remain comparable with the broader fund literature, including Ferreira et al. (2013), but this is acknowledged as a limitation.

Beyond the conditional/unconditional debate, factor alpha is only one way to measure performance. Cogneau and Hübner (2020) catalogue 147 distinct performance measures, and Kuosmanen (2007) proposes a non-parametric alternative that benchmarks funds without relying on any factor model at all. Alpha is retained here because it directly isolates risk-adjusted abnormal return — the quantity that matters most when comparing two management styles — but it captures only one dimension of a fund's performance.

3.4. Related Work: Previous Empirical Comparisons of the Two Approaches

A smaller and more recent body of literature places the two approaches in direct competition, comparing the realised performance of funds run by computerised models against that of funds run by human managers. The present study contributes directly to this literature, whose findings remain, to date, largely inconsistent.

Within the US equity mutual fund universe[17], Miguel and Chen (2021, 2025) provide two complementary contributions. In the first, they study the performance persistence of quantitative funds using the Lipper classification, which flags as quantitative, funds that buy and sell securities through a rule-based mathematical model. They find that whatever persistence exists is concentrated among poor performers, and that top-performing quantitative funds revert significantly more strongly than their human-managed peers — a pattern they interpret as evidence that quantitative funds exhibit less skill and flexibility. Their second paper turns to active management itself and documents that closet indexing[18] is widespread among quantitative funds: on average, 35% of the assets in active quantitative funds are managed by closet indexers throughout the sample period, rising to 50% by the end of 2019, against 24% on average for human-managed funds. Their active share[19] is also 5% lower on average than the one of non-quantitative peers. Net of fees, quantitative funds trail their benchmarks across every category of active management examined, and most heavily so among stock pickers. The study further complicates the common belief that quantitative funds are simply cheaper: while their total expense ratio is around 10 basis points lower on average, this advantage narrows substantially for high active-share strategies and is statistically indistinguishable from zero for stock

[17] Abis (2020) further shows, on a large sample of US equity funds classified through textual analysis of fund prospectuses, that quantitative funds tend to outperform discretionary peers in calm market conditions but give back that advantage during downturns, when the historical patterns embedded in their models no longer hold. She also documents that quantitative funds hold significantly more stocks than their discretionary counterparts, tend to favour stocks with greater information availability, charge lower fees on average, and are more prone to crowded trades — as many systematic funds draw on the same published signals and data sources. As a working paper, it is cited here in footnote form only, though much of the literature reviewed in this section builds on it.

[18] Closet indexing refers to the practice of funds that charge active management fees while maintaining a portfolio that closely tracks a passive benchmark, offering investors little genuine active exposure for the price they pay. (Miguel & Chen, 2025)

[19] Active Share measures the percentage of a fund's portfolio holdings that differ from those of its benchmark index, ranging from 0% — perfect replication — to 100% — no overlap whatsoever. Miguel and Chen (2025) use it as the central metric to assess how genuinely active quantitative funds are in practice.

pickers. Taken together, these two studies draw a picture in which systematic discipline has not, on average, translated into superior performance for mutual fund investors.

The narrower segment of AI-driven funds tells a different story. Studying fifteen AI-powered funds over 2017–2019, Chen and Ren (2022) confirm that these funds do not beat the market on a standalone basis, but show that once each fund is matched to similar human-managed peers, the AI funds outperform by close to 5.8% per year on a net basis. Much of this gap traces back to a structural cost advantage: AI-powered funds turn over their portfolios at a rate of roughly 31% per year, compared to 72% for their human-managed peers, and hold fewer stocks on average. Beyond cost, they also exhibit superior stock-selection ability and a measurably lower incidence of well-documented behavioural biases such as the disposition effect. The finding is encouraging for the systematic camp, but it rests on a very small and very recent sample, which limits how far it can be generalised.

Evidence from the hedge fund universe is equally divided. Chincarini (2014) compares quantitative and qualitative managers using the HFR database over 1994–2009 and finds that both groups earn positive risk-adjusted returns, with quantitative funds ahead by as much as 72 basis points per year on a ten-factor model — a difference he attributes to superior market-timing ability. Harvey et al. (2017) revisit the same setting over 1996–2014, covering a large set of HFR funds classified into systematic and discretionary strategies. For equity funds, they report near-identical risk-adjusted returns of 1.11% for systematic and 1.22% for discretionary managers; the appraisal ratio[20] in fact favours the systematic side slightly, at 0.35 against 0.25. For macro funds, systematic managers post a higher alpha of 4.85% against 1.57% for discretionary, though the gap narrows considerably once the higher volatility of systematic returns is taken into account. Their overall conclusion is that the two styles have delivered similar risk-adjusted performance, and that the reticence of many allocators to invest in systematic funds is not justified by the data — systematic managers were managing a disproportionately small share of assets (26%) relative to their share of the fund population (31%). Harvey et al. (2017) also show that discretionary funds load more heavily on well-known risk factors in aggregate, which means that a larger part of their raw returns reflects compensated risk exposure rather than genuine alpha. those two studies working on essentially the same database reach really different conclusions — Chincarini (2014) finding a 72-basis-point advantage for quantitative funds, Harvey et al. (2017) finding near-parity — illustrates how sensitive results are to the choice of model and classification method.

Moving beyond the hedge fund world, Zhang et al. (2023) examine nineteen big-data funds in China and reach a negative verdict: only seven of these funds outperform more than half of their matched

[20] The appraisal ratio is the ratio of a fund's risk-adjusted return (alpha) to the volatility of that same risk-adjusted return, making it the risk-adjusted analogue of the Sharpe ratio. Harvey et al. (2017) use it as their primary performance metric to compare systematic and discretionary hedge funds on a like-for-like basis.

traditional peers, and only three produce a statistically significant positive alpha differential. The big-data factor does not improve stock selection, and performance remains heavily dependent on individual manager skill rather than on the technology itself.

These studies contradict one another but they also cover slightly different side of the same question. They differ in the fund category of interest, in the technology being assessed, in geography, in sample length, in the way funds are classified as systematic, and in the benchmark model used — the Chincarini (2014) versus Harvey et al. (2017) comparison demonstrates how much the last two of these dimensions can matter. The study design adopted in the remainder of this master's thesis is intended to hold several of these dimensions constant simultaneously — a single classification rule applied to a consistent universe, a six-factor region-matched specification, a broad international sample of equity mutual funds, and a twenty-four-year window spanning multiple market regimes. The goal is to reduce the risk that observed performance differences reflect the particularities of the setting rather than real differences in management style.

4. Data and Methodology

This chapter lays out the empirical foundations of the study. Section 4.1 describes the data source and the construction of the sample. Section 4.1.5 defines every variable used in the analysis. Section 4.2 specifies the research design through which the two research questions are addressed. The empirical strategy follows the fund performance evaluation literature closely so that any difference detected between quantitative and discretionary funds is founded on an established methodological benchmark.

4.1. Data

4.1.1. Source and coverage

The data is from Refinitiv Lipper, accessed through LSEG Workspace (formerly distributed as Eikon). Lipper is one of the most comprehensive commercial databases for mutual fund research and has been used broadly in fund performance studies (Ferreira et al., 2013; Miguel & Chen, 2021, 2025). For each fund, the database provides monthly time series on: total return index (RI)[21], total net assets (TNA), and total expense ratio (TER) together with static data: launch and close dates, base currency, asset type, geographical focus[22], the Lipper strategy classification, share-class identifiers and the quantitative attribute.

The data extracted spans December 1999 to January 2025. The first and last month serve only to construct the returns, flows, and lagged controls of the first and last months of the analysis window: a monthly return requires two consecutive index levels, so the earliest observation serves as the lagged base for the first computable return. The samples of monthly fund returns therefore runs from January 2000[23] to December 2024, covering 300 months. The endpoint reflects the most recent reliable extract available at the time of data collection. All returns are expressed in US dollars, which removes the need for currency conversions across funds with different base currencies.

[21] The total return index shows the growth in value of a fund over a specified period (here 1 month), assuming that dividends are re-invested to purchase additional shares of the fund at the net asset value (NAV) on the dividend ex-date. (Lipper)

[22] According to Lipper the geographical focus classifies where a fund invests primarily. To be attributed to a region a fund must invest at least 70% of its capital in that specific region.

[23] Setting the start of the sample at the turn of the century follows Abis (2020): the 1998 amendment of disclosure that all funds were obligated to follow as of December 1999, added the principal investment strategy (PIS) section to funds' prospectuses. In this section, fund managers are required to explain how investments decisions are constructed giving a reliable base to classify funds by investment strategies.

4.1.2. Unit of observation: the primary share-class

All variables are measured at the share-class level, and for each fund only the primary share class[24] flagged by Lipper is retained. Encompassing multiple share classes into a single fund-level series would require complicated computation that the research question does not need: classes denominated in different currencies, carrying different fee schedules, and following different distribution policies coexist under one umbrella. The primary share class offers the cleanest representation of each fund. It also corresponds to the granularity at which investors allocate capital and at which fees are charged. The trade-off is that the results describe primary share classes rather than full umbrella fund, a point returned to in Section 4.2.4.

4.1.3. Identifying quantitative and discretionary funds

The classification of each fund as quantitative or discretionary forms the central explanatory variable of the study. A fund is labelled quantitative when its disclosed principal investment strategy (PIS) relies on systematic models, algorithmic signals, or model-driven decision rules. The classification rests on Lipper's own definition, which describes a quantitative fund as one "that uses primarily financial computer models and databases to determine the components of its portfolio". Every fund that does not meet this description is labelled discretionary. Borderline cases, funds that combine human judgement with a systematic component, are assigned to the quantitative category only when the systematic element is presented as the dominant driver of portfolio construction. This classification method follows Miguel and Chen (2021, 2025).

The classification is treated as static: each fund receives a single, time-fixed, label recorded in the most recent (end-of-period) data extract, so that a fund is considered quantitative or discretionary throughout the whole of its history in the panel. It does abstract from the small number of funds that may have genuinely shifted approach over the sample period. Because the classification ultimately rests on prospectus language, it could also carries some measurement noise, an issue acknowledged in the literature (Grobys et al., 2022; Harvey et al., 2017; Miguel & Chen, 2021, 2025) and revisited in Section 4.2.4.

The raw universe contains 3,235 funds, of which 246 are classified as quantitative which account for roughly 8% of the universe. In the main sample, 222 out of 3,031 funds are quantitative, representing 7.3%. This share is lower than the 20–30% reported by studies that classify funds through keyword

[24] The primary share class of a fund is the oldest retail share class available. If a fund has multiple share classes launched at the same time, the accumulating one is the primary share class. (Lipper)

counting of prospectus texts (Birru et al., 2024; Chincarini, 2014; Dyer et al., 2025). They are, however, consistent with Miguel and Chen (2021, 2025) results that use the same database and classification criteria. This gap is explained by the rule: Lipper's prospectus-based quantitative flag identifies only funds whose process is predominantly systematic and is therefore more conservative than a word-frequency approach.

4.1.4. Samples construction and data cleaning

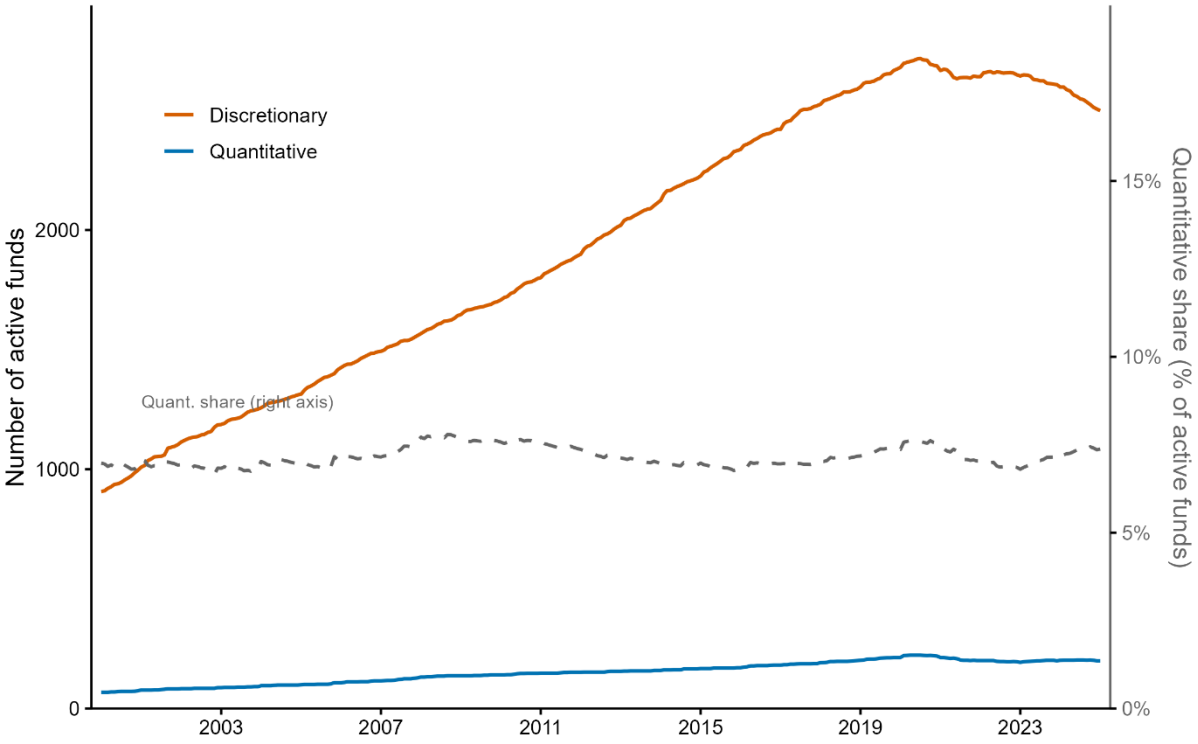
The starting universe consists of 3,235 US-domiciled[25], open-end, actively managed[26] equity mutual funds, of which 246 are quantitative and 2,989 are discretionary. Three cleaning operations are applied at the panel level before any sample is defined, so that the raw universe used for descriptive figures, the main sample used for the core analysis, and the appendix robustness samples all rest on the same cleaned data.

(1) Each fund's time series is trimmed to its operating life. Observations dated before the launch or after the closure of a fund are removed. Lipper sometimes carries the last value of a variable forward after closure or attaches the pre-registration track records of incubated portfolios to the fund's history. Because asset managers typically bring to market only those incubated portfolios whose early results were favourable, retaining these observations would import a well-known upward bias, the incubation bias, into the performance estimates (Evans, 2010). This life-span trimming removes 20,942 observations across 593 funds. (2) Net returns, gross returns, and net flows are winsorised at the 1st and 99th percentiles, consistent with the literature ((Dyer et al., 2025; Ferreira et al., 2013; Miguel & Chen, 2021, 2025). Databases are known to carry false extreme values on returns, size and flow coming from recording errors, stale prices, or mechanical artefacts which would drive the estimates of our models. (3) The total expense ratio is left exactly as Lipper reports it; a month with a missing TER produces a missing gross return for that month but does not affect the net return or any other variable.

The raw universe then consists of every fund with at least one observation over the analysis window of January 2000 to December 2024. This leaves 3,231 unique funds (4 funds are dropped from the starting universe because they launched in December 2024), 246 quantitative and 2,985 discretionary. The raw universe is used only to describe the evolution of the fund population over time. **Figure 1** shows the resulting fund repartition and the quantitative share of the total active population.

[25] The quantitative attribute is only available for US-domiciled funds certainly due to the fact that they are all required to report the same type of prospectus. (Miguel & Chen, 2021, 2025)
[26] Index-tracking funds, exchange traded funds and Funds of funds are excluded of the extract. It is not relevant to include them as they are generally passive investment vehicles.

Figure 1: Evolution of the number of funds in the sample



The **main sample** is obtained from this raw universe by imposing three additional filters, following Miguel and Chen (2021). First, a minimum of 36 monthly return observations, in order to ensure sufficient time-series data to estimate the six-factor alpha (Carhart, 1997; Fama & French, 2010; Grinblatt & Titman, 1994; Miguel & Chen, 2021) Second, availability of all control variables: log TNA, TER, net flow, and age (Miguel & Chen, 2021; Chen et al., 2004; Ferreira et al., 2013). Third, a valid geographical focus mapping to one of the seven regional factor sets to guarantee regional factor matching. These conditions drop 203 funds, most of them recently launched: the median dropout has fewer than eight return-months. The main sample therefore contains **3,028 unique funds**, 222 quantitative and 2,806 discretionary, and **626,743 fund-month observations** (an average of 207 observation per fund, well above the 36-month estimation floor). Its regional composition is given in **Table 1** and is dominated by United States funds (1,896), followed by Global ex US (463), Global (414), and Global Emerging Markets (224). Europe, Japan, and Asia Pacific ex Japan each contribute around ten funds, and the three together account for only 1.2% of the total fund-months in the sample, confirming their negligible weight in the comparison. Surviving and liquidated funds are both retained, since conditioning on survival would mechanically inflate measured performance what is defined as

the survivorship bias and documented by Carhart (1997), and Fama and French (2010). This is the reason why most of the literature on fund performance use survivorship-bias-free databases.

Table 1: Sample composition by style and geographical focus

	Quantitative		**Discretionary**		**Total**	
Geographical focus	**Funds**	**Fund-months**	**Funds**	**Fund-months**	**Funds**	**Fund-months**
United States of America	145	31,479	1,751	392,400	1,896	423,879
Global	26	4,109	388	66,596	414	70,705
Global ex US	34	6,354	429	82,054	463	88,408
Global Emerging Markets	17	2,844	207	33,331	224	36,175
Europe	0	0	11	2,829	11	2,829
Japan	0	0	11	2,555	11	2,555
Asia Pacific ex Japan	0	0	9	2,192	9	2,192
Total	**222**	**44,786**	**2,806**	**581,957**	**3,028**	**626,743**

Notes: Analysis sample January 2000–December 2024 (300 months). A fund is included if it has at least 36 non-missing monthly returns, data on all control variables and a valid geographical focus. Fund-months count non-missing return observations.

Two **robustness samples** are constructed for the appendix, each modifying a single element of the main sample. The first lowers the return-history threshold from 36 to 24 months (Ferreira et al., 2013; Zhang et al., 2023), reintroducing 93 shorter-lived funds for a total of 3,121 funds (226 quantitative representing of the sample 7.2% and 2895 discretionary). The second raises the threshold to 60 months, restricting the sample to 2,847 (211 quantitative representing 7.4% of the sample and 2636 discretionary) funds with a longer track record. Both retain the control-availability and geographical-focus requirements of the main sample.

4.1.5. Variable description

The empirical analysis is based on three groups of variables: performance measures, which serve as dependent variables; the style indicator (quant dummy), which is the key explanatory variable; and a set of fund-level controls. Most variables come directly from Lipper; the remainder are standard transformations of Lipper data or are estimated from the factor regressions described below. Returns, flows, and expense ratios are expressed in percentage terms, while TNA is reported in USD millions. **Table 2** consolidates the definition, construction, and source of every variable used in the study.

Table 2: Variables description

Variable	Definition
Net return	Monthly total return of fund *i* in month *t*, in US dollars and net of fees, computed from the Lipper total return index (Lipper): $$R_{i,t} = \frac{RI_{i,t}}{RI_{i,t-1}} - 1$$
Gross return	Monthly return before fees of fund *i* in month *t*, approximated by adding back the monthly expense ratio (own computation from Lipper): $$R_{i,t}^{gross} \approx R_{i,t} + TER_{i,t}^{m}$$
Excess return	Net (or gross) return in excess of the one-month risk-free rate of the fund's matched region. Risk-free rates are taken from the Kenneth R. French Data Library[27].
Net alpha (%)	Intercept of the six-factors time-series regression (Equation 1) on net excess returns, estimated fund by fund, annualised and expressed in percentage per year (own estimation).
Gross alpha	As above, estimated on gross excess returns (own estimation).
Quantitative	Dummy variable that takes the value of 1 if the fund is classified as quantitative, and 0 if discretionary. The classification is time-invariant: one label per fund, retrieved at the end of the sample period (December 2024) (Lipper).
TNA	Total net assets of the fund share class, in millions of US dollars (Lipper).

[27] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

Variable	Definition
Log TNA	Natural logarithm of TNA, used to address the right-skew of the size distribution (own computation).
TER	Total annual expenses as a fraction of net assets (in percentage per year) (Lipper). The monthly expense ratio used to estimate gross returns is: $$TER_{i,t}^{m} = \frac{TER_{i,t}}{12}$$
Flow	Percentage growth in TNA in each month, net of internal growth. Fund flow for fund *i* at period *t* is computed as: $$Flow_{i,t} = \frac{TNA_{i,t} - TNA_{i,t-1}(1 + R_{i,t})}{TNA_{i,t-1}}$$ where $TNA_{i,t}$ is the total net asset value of fund *i* at the end of month *t*, and $R_{i,t}$ is its net return. (own computation from Lipper).
Age	Number of years since the fund launch date (Lipper): $$Age_{i,t} = \frac{date_t - launch\ date_i}{365.25}$$
Geographical focus	Lipper's stated investment-region category, used to match each fund to its regional factor set and to define region fixed effects (Lipper).
MKT, SMB, HML, RMW, CMA, MOM	Monthly returns on the market, size, value, profitability, investment and momentum factors, matched to each fund's geographical focus (Kenneth R. French Data Library).

Note: The net flow follows the literature, which derive it from the change in net asset value, net of investment performance (Berk & Green, 2004; Miguel & Chen, 2021).

4.2. Methodology

The empirical strategy is decomposed in two stages. The first stage produces a single risk-adjusted performance measure for each fund. The second stage compares those measures across the two management styles, first without and then with controls. The design maps directly onto the two research questions: the style coefficient estimated in the second stage answers **RQ1**, on whether quantitative funds outperform discretionary funds on a risk-adjusted basis, while the comparison of net and gross performance, the role of the controls, and the regional effect respond to **RQ2**.

4.2.1. Risk adjustment and regional factor matching

Performance is measured with the Fama and French (2015) five-factor model augmented with a momentum factor (Carhart, 1997; Jegadeesh & Titman, 1993), giving the six-factor specification (FF6) introduced in section 3.3.3. The purpose, tracing back to Jensen (1968), is to separate risk-adjusted performance from returns that merely compensate exposure to well-known systematic risks: a fund can post high gross returns by loading on the market, small caps, value stocks, or past winners, and only the portion of its return unexplained by those exposures is interpreted as alpha. The six-factor model is adopted as a pragmatic and widely accepted benchmark rather than an absolute definition of expected returns. The proliferation of candidate factors and the multiple-testing concerns raised by Harvey et al. (2016) warns against reading the estimated alphas as anything more than performance relative to a standard, transparent risk model.

Because the funds in the sample specialise in different regions, applying a single set of factors to every fund would misrepresent the risk exposures of internationally focused funds. Each fund is therefore matched, through its Lipper geographical focus, to the corresponding set of regional factors drawn from the Kenneth R. French Data Library. The mapping pairs US-focused funds with the North America factors, Global funds with the Developed-market factors, Global ex US funds with the Developed ex US factors, Global Emerging Markets funds with the Emerging-market factors, and the European, Japanese, and Asia-Pacific ex Japan funds with their respective regional factors. Excess returns are formed against the risk-free rate of the matched region, and fund months returns are aligned exactly with factor months. Regional matching matters because, although factor premia exists internationally, they differ substantially across regions, and local factor models outperform global ones in describing local returns (Asness et al., 2013; Fama & French, 2010; Ferreira et al., 2013).

4.2.2. First stage: fund-level alpha estimation

For each fund, monthly excess returns are regressed on the six matched factors by ordinary least squares regression:

Equation 1: The six-factors alpha estimation (FF6)

$$R_{i,t} - RF_{i,t} = \alpha_i + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_5 CMA_t + \beta_6 MOM_t + \varepsilon_{i,t}$$

where $R_{i,t}$ is the return in US dollars, net or gross of fees, of fund *i* in month *t*; $RF_{i,t}$ is the monthly risk free rate matched to the fund's geographical focus; MKT_t is the excess return on the market portfolio; SMB_t(small minus big) is the return spread between small and large capitalisation stocks; HML_t(high minus low) is the return spread between high and low book-to-market stocks; RMW_t(robust minus weak) is the return spread between firms with robust and weak operating profitability; CMA_t(conservative minus aggressive) is the return spread between firms with conservative and aggressive investment policies; and $MOM_t$is the return spread between past twelve-month winners and losers. ε_i is the residual. **Equation 1** is estimated separately for each fund, yielding a fund-specific intercept α_i that captures the fund's average monthly abnormal return above what would be predicted by exposure to the six systematic factors. This intercept is then annualised by multiplying by twelve and serves as the primary performance measure in the cross-sectional analysis. A parallel set of regressions is run on gross-of-fee returns, obtained by adding back the monthly expense ratio, to distinguish between value destruction or creation before and after costs. The gross regression is subject to the same 36-month minimum as the net regression; funds with fewer than 36 months of non-missing TER data receive no gross alpha estimate, which exclude 94 funds from the sample.

4.2.3. Second stage: comparing the two styles

The fund-level alphas are then compared across styles through two complementary procedures, one univariate and one multivariate.

The univariate procedure is a series of Welch two-sample t-tests, which compare the mean of each variable — net alpha, gross alpha, TER, TNA, age, and net flow — between quantitative and discretionary funds. A two-sample t-test asks whether the difference between two group means is large enough, relative to its sampling variability, to be unlikely to have arisen by chance. The Welch (1947) variant is preferred to the classical Student formulation because it does not assume that the two groups share a common variance: it estimates each group's variance separately and adjusts the

degrees of freedom accordingly. This matters here because the quantitative and discretionary subsamples are of very unequal size — quantitative funds make up roughly 7% of the universe — and differ in their dispersion of size, fees, and flows, precisely the conditions under which the equal-variance assumption fails and the classical test understates the standard error of the difference. The Welch test therefore provides an unconditional representation of the style gap that is robust to this imbalance.

The multivariate procedure is a cross-sectional regression estimated with one observation per fund:

Equation 2:Cross-sectional style regression

$$\alpha_i = \gamma_0 + \gamma_1 \mathrm{Quant}_i + \gamma_2 X_i + \sum_r \delta_r \mathrm{Region}_{r,i} + \varepsilon_i$$

where α_i is the fund's annualised FF6 alpha in percent per year — either net or gross of fees depending on the specification — and Quant_i is a binary indicator equal to one if fund *i* is classified as quantitative and zero otherwise. Quant_i is the vector of fund-level controls — log TNA, age, TER, and net flow, each averaged over the fund's life. $\mathrm{Region}_{r,i}$ is a set of region fixed-effect indicators, equal to one if fund *i* invests primarily in region *r*, which absorb systematic performance differences across investment regions and, with them, the uneven geographical distribution of quantitative funds across the sample. ε_i is the residual. The regression is estimated by ordinary least squares with one observation per fund, and standard errors are heteroskedasticity-robust (White, 1980). The coefficient of interest is γ1: it measures the average risk-adjusted performance difference between quantitative and discretionary funds, holding size, age, fees, flows, and region constant. A positive and significant γ1 indicates that quantitative funds outperform; a coefficient close to zero suggests that observable characteristics and factor exposures account for the cross-style difference; a negative coefficient points to quantitative underperformance. This estimate is the direct answer to **RQ1**.

The variance in such a cross-section is unlikely to be constant across funds: the dispersion of alpha is mechanically larger for funds estimated over shorter histories or holding more concentrated portfolios, so the homoskedasticity assumption of ordinary least squares is implausible here. Under heteroskedasticity, the OLS coefficients remain unbiased but their conventional standard errors are inconsistent, which distorts the t-statistics and can make a coefficient appear significant when it is not. The standard errors reported throughout the second stage are therefore the heteroskedasticity-consistent estimator of White (1980), in its HC1 finite-sample form, which delivers valid inference

without imposing any structure on the form of the heteroskedasticity. Using a robust covariance estimator rather than the classical one is the norm in the empirical fund-performance literature: cross-sectional and panel studies of fund return. It corrects the standard errors for heteroskedasticity, autocorrelation, or cross-sectional dependence (Chincarini, 2014; Dyer et al., 2025; Harvey et al., 2017; Miguel & Chen, 2021; Yan, 2008).

The regression is estimated as a sequence of increasingly saturated specifications: the style indicator alone, then with controls, then with controls and region fixed effects, and finally the fully specified model re-estimated on gross alpha. Exploring the shift from the net to the gross specification isolates whether any style gap originates in fees or in pre-fee performance, which is central to **RQ2**. Including the controls is as important: quantitative and discretionary funds differ in size, age, and cost, which all relate to performance (Berk & Green, 2004; Ferreira et al., 2013; Pástor et al., 2015; Yan, 2008), so omitting these characteristics would bias the estimated style coefficient.

Because the controls correlate with the style indicator, the second stage is paired with a multicollinearity diagnostic: the pairwise correlation matrix of the regressors and the variance inflation factors (VIF)[28] of the fully specified model are inspected to confirm that the estimated style coefficient is not destabilised by collinearity. These diagnostics are reported in the appendix.

4.2.4. Robustness and design choices

All results in the main body originate from the main sample, whose construction is described in Section 4.1.4. The 36-month return-history requirement that anchors this sample ensures every first-stage alpha is estimated with enough precision to be meaningful, but it mechanically excludes short-lived funds. Two robustness samples address this. The first lowers the threshold to 24 months (Ferreira et al., 2013; Zhang et al., 2023) reintroducing the shorter-lived funds whose exclusion is the most plausible source of survivorship bias in the main estimates (Fama & French, 2010). The second raises it to 60 months, restricting attention to well-established funds. The full second stage is re-estimated on both, and the resulting style coefficients are reported in the appendix. The same logic applies to the choice of dependent variable and to the inclusion of control variables and region fixed effects.

[28] The variance inflation factor for regressor j is $VIF_j = 1/(1 - R^2_j)$, where R^2_j is the coefficient of determination from a regression of j on all remaining regressors. It measures how much the variance of the estimated coefficient on j is inflated relative to what it would be if j were orthogonal to the other regressors. A VIF of one indicates complete absence of linear dependence; the factor rises as collinearity increases

Beyond these robustness checks, three design choices deserve explicit acknowledgement:

The first concerns the unit of observation. Each fund enters the sample through its primary share class rather than as a consolidated umbrella fund. This avoids a genuine problem: aggregating across share classes denominated in different currencies, carrying different fee structures, and sometimes launched years apart requires manipulation that are difficult to realise in the database and easy to get wrong. The cost is a weakened representativeness — the results describe primary share class performance, and a different share class of the same fund might show a different alpha if its fee load or other characteristics differ.

The second concerns factor matching. Each fund is assigned to a regional factor set based on its stated geographical focus in Lipper. In practice, a fund labelled as European may hold meaningful positions in the United States or Asia, particularly during periods of market shifts. The matched factors then mismeasure the fund's true risk exposures, and some of what appears as alpha may in fact be uncompensated exposure to factors outside the matched region. This is a standard limitation of prospectus-based region assignment and cannot be resolved without holdings data at monthly frequency.

The third, discussed in Section 4.1.3, is that the quantitative classification is static. A fund classified as quantitative (discretionary) in 2024 keeps that label in 2001, even if it progressively adopted quantitative tools over the period of interest. Any performance difference attributed to style therefore reflects the label as assigned to a fund in December 2024 rather than the actual investment process at each point in time. Given that the boundary between the two styles has blurred over the sample period, this likely compresses the estimated style gap rather than inflating it.

5. Empirical Analysis and Results

The two-stage design set out in the methodology 4.2 is taken to the data in this chapter. Section 5.1 describes the observable characteristics of the two groups of funds. Section 5.2 gives a first comparison of the two styles through differences in means. Section 5.3 checks the regressors for collinearity, Section 5.4 estimates the cross-sectional regressions **Equation 2** and Section 5.5 re-estimates them on the two robustness checks.

5.1. Sample description and descriptive statistics

Before asking whether style affects performance, it is useful to characterise each group, since the differences between them are what the regression of Section 5.4 will hold constant. The regional breakdown in **Table 1** already points to one feature that shapes the comparison. Quantitative funds appear in only four of the seven regions: the United States (145 funds) and the three global markets covering developed markets, developed ex-US stocks, and emerging markets. None of the 222 quantitative funds invest primarily in Europe, Japan, or Asia-Pacific ex Japan. Systematic strategies generate returns by spreading many small, model-driven bets across a wide range of securities; a single-country universe constrains the investable universe in a way that a broad global or US-wide universe does not (Chincarini, 2014; Miguel & Chen, 2025). This concentration is the reason region fixed effects enter the second stage. Left uncontrolled, the style coefficient would absorb part of the average performance of the regions in which quantitative funds happen to cluster.

Table 3 summarises the funds characteristics, averaging each variable over the fund's life. The P10 and P90 columns show that both groups are far from homogeneous.

Table 3: Descriptive statistics for quantitative and discretionary funds

	Quantitative (N = 222)				Discretionary (N = 2,806)			
	Mean	**Med.**	**P10**	**P90**	**Mean**	**Med.**	**P10**	**P90**
Monthly net return (% p.m.)	0.76	0.76	0.41	1.08	0.77	0.78	0.43	1.10
	(0.29)				(0.30)			
Monthly gross return (% p.m.)	0.84	0.86	0.50	1.18	0.87	0.88	0.54	1.20
	(0.30)				(0.31)			
TNA (USD millions)	499	148	11	1,133	1,111	213	11	2,205
	(1,271)				(3,950)			
TER (% p.a.)	0.98	1.00	0.45	1.42	1.09	1.07	0.65	1.52
	(0.40)				(0.40)			
Age (years)	11.0	9.5	2.9	20.7	12.6	10.0	3.1	24.8
	(8.8)				(11.0)			
Net flow (% p.m.)	0.94	0.57	-0.80	3.12	0.91	0.52	-0.71	3.01
	(1.71)				(1.68)			

Notes: Standard deviations in parentheses. All variables are averaged over each fund's life. P10 and P90 are the 10th and 90th percentiles. Monthly gross return is approximated as net return plus one-twelfth of the annual TER. TNA is total net assets in USD millions. TER is the total expense ratio in % per year. Net flow is in % per month.

The profiles that emerge from **Table 3** is consistent with the literature[29] (Dyer et al., 2025; Miguel & Chen, 2021, 2025). Quantitative funds are smaller: their mean TNA is USD 499 million against USD 1,111 million for discretionary funds, the median gap is USD 148 million versus USD 213 million, a better indicator on the typical fund since a handful of very large discretionary funds inflates the mean. Quantitative funds are also younger by one to two years and carry a lower expense ratio, 0.98% against 1.09% per year. Net flows are indistinguishable, near 0.9% per month for both groups, and the bottom decile of funds in each style exhibits negative flows, which shows the sample retains funds under redemption pressure and not just growing ones. Raw monthly net returns are near-identical, 0.76% for quantitative funds and 0.77% for discretionary funds; the gross returns are equally close at 0.84% versus 0.87%, the gap tracking the TER difference almost exactly. Before any risk adjustment the two styles look alike on unadjusted returns; any performance difference must surface in the alpha.

[29] Abis (2020) observe similar results.

The size gap is worth keeping in mind for what follows. Quantitative funds are smaller, younger, and cheaper than their discretionary peers, and all three characteristics relate to alpha in this sample: size through the scale dynamics documented Yan (2008), and Pastor et al. (2015); fees through the direct drag on net returns; age through the selection effects built into a minimum-history sample. A direct comparison of average alpha across the two styles would therefore mix the effect of being quantitative with the effect of being smaller, cheaper, and younger. **Table 4** makes that comparison anyway as a starting point; Section 5.4 then holds those differences constant.

5.2. Univariate comparison: differences in means

Table 4 reports Welch two-sample t-tests comparing quantitative and discretionary funds on each main variable. The interpretation throughout is suggestive rather than decisive: the test ignores the characteristic differences that **Table 3** already shows to be significant, so it cannot isolate the effect of style from the effect of size, fees, age or flows.

Table 4: Differences in means between quantitative and discretionary funds

Variable	Quant (mean)	Discretionary (mean)	Difference (D−Q)	t-statistic	p-value
Alpha, net (% p.a.)	-0.48	0.21	0.68***	4.63	0.000
Alpha, gross (% p.a.)	0.50	1.36	0.86***	6.03	0.000
Monthly net return (% p.m.)	0.76	0.77	0.01	0.68	0.500
Monthly gross return (% p.m.)	0.84	0.87	0.03	1.33	0.184
TNA (USD millions)	499	1,111	612***	5.40	0.000
TER (% p.a.)	0.98	1.09	0.11***	3.82	0.000
Age (years)	11.00	12.60	1.60**	2.58	0.010
Net flow (% p.m.)	0.94	0.91	-0.03	-0.29	0.773

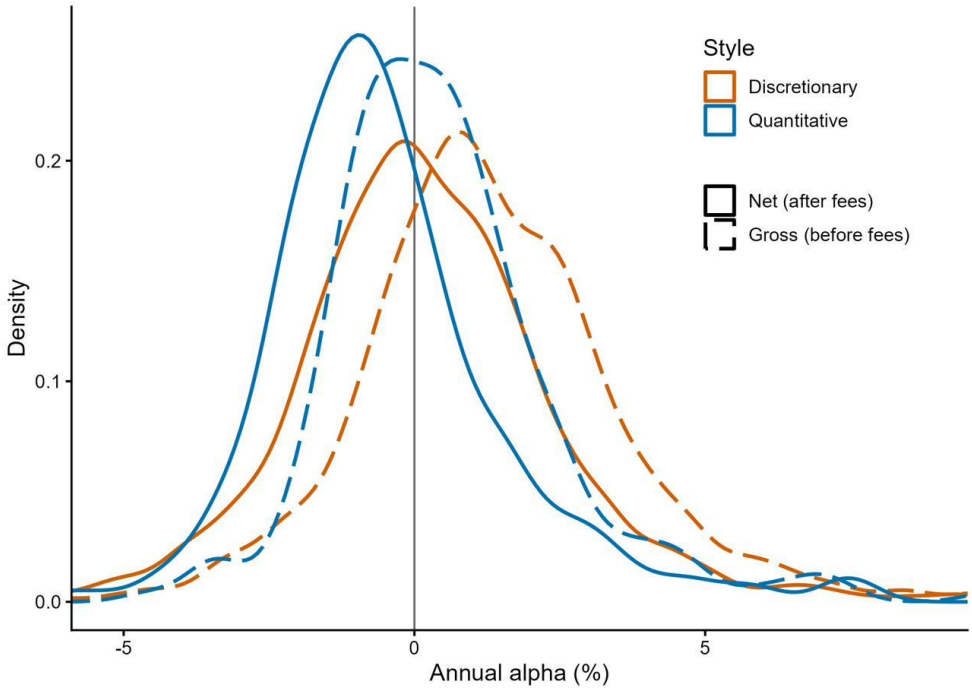
Notes: Welch two-sample t-test. Difference is Discretionary minus Quantitative. *** p<0.01, ** p<0.05, * p<0.10. Alpha is the annualised FF6 intercept in % per year. Monthly gross return is approximated as net return plus one-twelfth of the annual TER. TNA is total net assets in USD millions. TER is the total expense ratio in % per year. Age is in years. Net flow is in % per month. Sample: main sample (minimum 36 months).

The performance gap is clear in both alpha measures. Quantitative funds earn on average 0.48% per year net of fees, discretionary funds earn 0.21%, a difference of 0.68% significant at the 1% level. Gross of fees the gap widens to 0.86 percentage points: 0.50% for quantitative funds versus 1.36% for

discretionary funds. The gross gap exceeds the net gap because quantitative funds charge lower fees (0.98% versus 1.09%), which partially offsets their pre-fee underperformance but does not eliminate it. Raw monthly returns, by contrast, are not significantly different between the two groups, whether net or gross. The performance gap is therefore a risk-adjusted phenomenon: it only surfaces once common factor exposures are removed, and not in the unadjusted return series. Among the characteristics, the size gap is the largest in absolute terms: discretionary funds are on average USD 612 million larger, significant at the 1% level. The fee and age gaps are smaller but also significant: discretionary funds charge 0.11 percentage points more per year and are on average 1.6 years older. Net flows are the exception, with no meaningful difference between the two groups, which could suggest that investor demand has not yet priced in whatever performance difference exists.

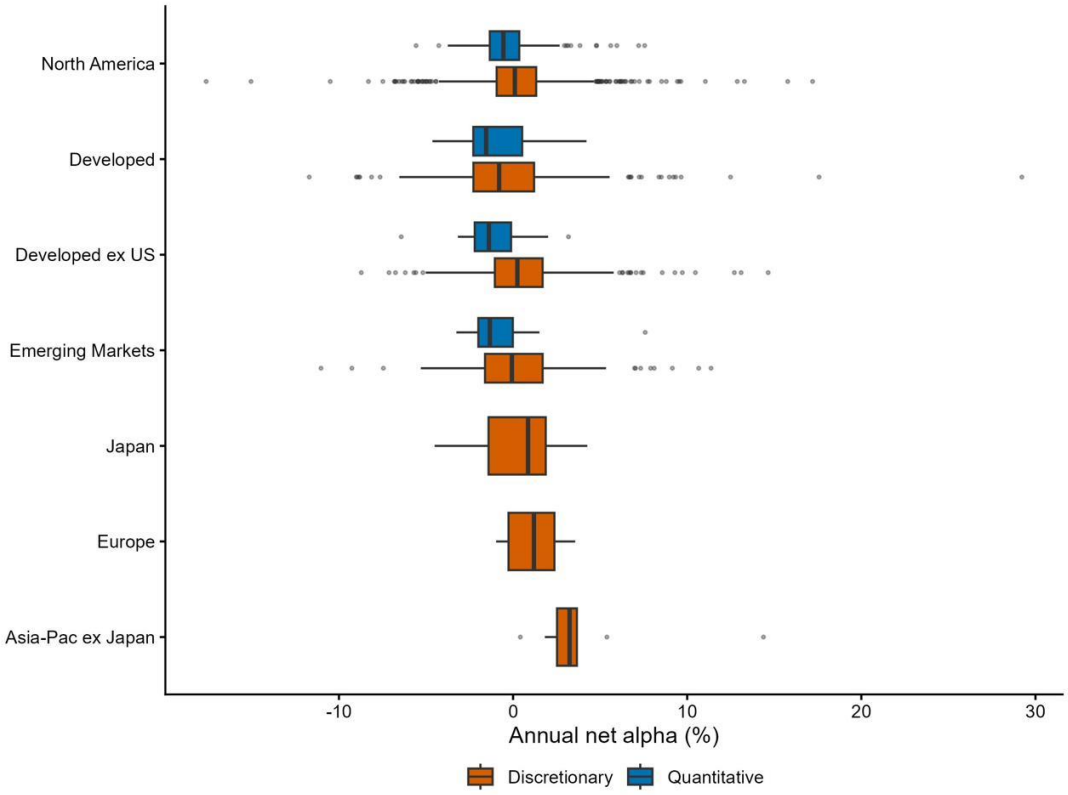
Figure 2 and **Figure 3** put these means in distributional context. **Figure 2** shows the smoothed distribution of annual alpha for each style, net and gross of fees. The net distributions of both styles sit close to zero, with the quantitative curve to the left of the discretionary one and carrying a lighter right tail; the shift from net to gross moves each curve to the right by roughly the expense ratio. This pattern matches the central finding of Fama and French (2010): gross of fees the typical fund approximately tracks its risk benchmark, and fees push the net distribution below zero on average.

Figure 2: Distribution of annual alpha by style



Note: Smoothed distribution of the annualised fund-level FF6 alpha, one observation per fund. Solid lines show net-of-fee alpha, dashed lines gross-of-fee alpha. The vertical line marks zero. Axis limits trim the outer 1% of each tail for readability

Figure 3: Annual net FF6 alpha by region and style



Note: Each box shows the median and middle 50% of the net alpha distribution for that group and region. Individual points are funds with unusually high or low alpha.

Figure 3 represents the distribution of the net alpha across regions. In the four regions where both styles operate, the quantitative boxes sit at or just below the discretionary boxes and are tighter, with fewer funds in the extreme tails. The discretionary funds have long right tails, mostly in North America and Developed markets. The three discretionary only regions post positive median net alphas, but each is constructed on about ten funds and carries almost no weight in the pooled comparison.

5.3. Correlation structure and multicollinearity

Before estimating the regression (**Equation 2**), two checks confirm that the five right-hand-side variables can sit in the model without distorting one another. The first is the pairwise correlation matrix in **Table 5**, the second is the set of variance inflation factors (VIF) in **Table 6**.

Table 5: Correlation matrix of the regressors

Variable	Alpha, net (%)	Alpha, gross (%)	Quant (=1)	Log TNA	Age (years)	TER (%)
Alpha, net (%)	1					
Alpha, gross (%)	0.943	1				
Quant (=1)	−0.069	−0.089	1			
Log TNA	0.151	0.094	−0.046	1		
Age (years)	0.009	0.002	−0.039	0.463	1	
TER (%)	−0.029	0.079	−0.070	−0.270	0.058	1
Net flow (%)	0.115	0.055	0.005	−0.253	−0.467	−0.144

Notes: Lower-triangular Pearson correlation matrix. *** p<0.01, ** p<0.05, * p<0.10. N = 2,934 funds. Quant (=1) is the style indicator. Log TNA is the log of total net assets in USD millions. TER is the total expense ratio in % per year. Net flow is in % per month.

The matrix points to two features relevant to the regression. The style indicator correlates weakly with every control: its strongest link is with the expense ratio (−0.070), and the rest fall below 0.05 in absolute value. The style coefficient estimated in Section 5.4 therefore measures style, not a control variable it happens to track. The larger correlations sit among the controls themselves, between size and age (0.463) and between age and net flow (−0.467), and even these stay moderate. The bivariate links between the controls and net alpha follow the patterns already set out in Section 5.1: size and net flow correlate positively with alpha (0.151 and 0.115), while age and the expense ratio show no net relationship. Whether each survives once the others are held constant is the question the regression answers.

Table 6: Variance inflation factors (VIF)

Regressor	VIF
Quantitative (=1)	1.011
Log TNA	1.454
Age (years)	1.558
TER (%)	1.168
Net flow (%)	1.316

Notes: VIF from the OLS of net alpha on all regressors without region fixed effects. A VIF above 10 indicates severe multicollinearity; below 5 is unproblematic. Sample: main sample (minimum 36 months).

The variance inflation factors point the same way. Every value lies between 1.01 and 1.56, well under the conventional threshold of five above which collinearity starts to inflate standard errors. Age carries the highest factor at 1.56, in line with its correlation with size and flow, but stays far from a level that would alter the individual estimates. The coefficients in Section 5.4 can be interpreted independently.

5.4. Cross-sectional regressions of fund alpha on style and controls

Table 7 reports the cross-sectional regressions **Equation 2** results that give a definitive answer to **RQ1**. Each column is an OLS on one observation per fund, with heteroskedasticity-robust (HC1) standard errors. The dependent variable is the annualised FF6 alpha from the first stage. Columns (1) to (3) use net-of-fee alpha and build from the raw style gap to the full specification; column (4) repeats the full specification on gross-of-fee alpha.

Table 7: Cross-sectional regressions of fund alpha on style and controls

	Dependent variable: FF6 alpha (% p.a.)			
	Net of fees	**Net of fees**	**Net of fees**	**Gross of fees**
	(1)	**(2)**	**(3)**	**(4)**
Quantitative (=1)	**−0.683*** **	**−0.567*** **	**−0.565*** **	**−0.688*** **
	(0.147)	(0.147)	(0.146)	(0.141)
Log TNA		0.251***	0.244***	0.206***
		(0.031)	(0.031)	(0.028)
Age (years)		−0.004	−0.005	−0.010**
		(0.005)	(0.005)	(0.004)
TER (%)		0.322*	0.325*	0.913***
		(0.180)	(0.185)	(0.159)
Net flow (%)		0.258***	0.264***	0.176***
		(0.050)	(0.050)	(0.047)
Controls	No	Yes	Yes	Yes
Region FE	No	No	Yes	Yes
N	3,028	3,028	3,028	2,934
Adj. R^2	0.004	0.052	0.064	0.054

*Notes: Heteroskedasticity-robust standard errors (HC1) in parentheses. *** p<0.01, ** p<0.05, * p<0.10. One observation per fund. Column (4) has 94 fewer observations than columns (1)–(3) because the gross regression applies the same 36-month minimum as the net regression; funds with fewer than 36 months of non-missing TER data receive no gross alpha estimate. Dependent variable: annualised FF6 alpha in % per year. Controls: Log TNA (log USD millions), Age (years), TER (%), and net flow (% per month), each averaged over the fund's life. Sample: main sample (minimum 36 months).*

Column (1) regresses net alpha on the style dummy alone. Quantitative funds earn 0.68 percentage points less per year than discretionary funds, significant at the 1% level. The coefficient reproduces the univariate gap of Section 5.2, as it should, since with no other regressor the dummy returns the difference in group means.

Column (2) adds the four fund-level controls. The style coefficient moves from −0.683 to −0.567 and keeps its significance at the 1% level. Adding the four controls reduces the style gap by 0.12 percentage

points, from −0.683 to −0.567, suggesting that a small part of the raw underperformance reflects the characteristics of quantitative funds rather than their investment approach. Among the controls, size and net flow coefficients are positive and significant: a fund twice as large in assets earns around 0.17[30] percentage points more alpha per year, and a one point increase in net flow adds 0.26 percentage points, in line with the flow-performance relation of Berk and Green (2004), Age is flat, and the expense ratio coefficient is positively but is significant only at the 10% level for the net of fees specification.

Column (3) adds region fixed effects. Quantitative funds concentrate in the United States and the three global regions (section 5.1), so the style coefficient could in principle absorb the average performance of those regions. Holding region constant shows that the coefficient barely moves, from −0.567 to −0.565. The quantitative penalty holds within regions, not only across them.

The estimate stays close to −0.57 percentage points per year across the three net columns. Taken together they answer **RQ1**: after risk adjustment and observable controls, the average quantitative fund earns lower net alpha than the average discretionary fund, and neither fund characteristics nor regional composition removes the gap. Both styles earn average net alpha close to zero, so neither group generates meaningful abnormal returns in absolute terms. Within that narrow range, however, a gap of 0.57 percentage points separates them, and the estimate is significant at the 1% level: the style coefficient carries a t-statistic above 3 across all three net columns.

Column (4) turns to gross alpha. Gross returns are constructed by adding back the monthly TER to net returns, so the only difference between the two first-stage regressions is a constant added to the dependent variable each month. That constant shifts the intercept by the same amount, leaving the factor loadings unchanged, which is why the style gap widens from −0.565 net to −0.688 gross by roughly the 0.11 percentage point average fee difference between the two groups (Section 5.2). Quantitative funds benefit from lower costs, but the fee advantage is not large enough to offset their wider pre-fee underperformance. The expense ratio coefficient confirms this: marginally significant net (0.32), strongly significant gross (0.91), the difference between the two reflecting the fee added back

The adjusted R^2 stays low throughout, between 0.05 and 0.06, meaning the model explains only 5 to 6% of the variation in alpha across funds. This is expected: once market, size, value, profitability, investment, and momentum exposures are stripped out by the first-stage regression, what remains is largely fund-specific noise that observable characteristics like size, fees, or investment style cannot systematically predict. Fama and French (2010) document the same pattern. The low R^2 does not

[30] Since log(2) ≈ 0.693. Multiplying the coefficient by 0.693 gives 0.251 × 0.693 ≈ 0.17 percentage points.

undermine the style coefficient: with 3,028 funds, a precisely estimated gap of −0.57 percentage points can coexist with a model that explains little of the total variation, because most of that variation is idiosyncratic rather than systematic[31].

5.5. Robustness

Section 4.1.4 defines two samples that differ from the 36-month return-history floor of the main analysis, the lower at 24 months and the upper at 60. Re-estimating columns (1) to (4) of **Table 7** on each does not change the answer to RQ1. The net style coefficient from the fully specified net model (column (3)) is −0.61 at the 24-month floor and −0.69 at the 60-month floor, against −0.57 in the main sample, and the gross coefficient (column (4)) lies between −0.65 and −0.78 across the same range. Every estimate keeps its negative sign and its significance at the 1% level.

Read in threshold order, the coefficients answer the survivorship concern raised in Section 4.1.4. If short-lived quantitative funds excluded by the history floor were the ones pulling the average down (Fama & French, 2010), then restoring them at 24 months should widen the gap, and keeping only established funds at 60 months should narrow it. The estimates move the other way. The gap holds when the short-lived funds re-enter at 24 months and reaches its widest point among the 60-month funds. The shortfall does not come from the short-lived tail; the funds with the longest records show the largest gap.

The control coefficients keep the pattern of **Table 7** in both samples. Size and net flow coefficients stay positive and significant at the 1% level, and the age coefficient stays close to zero. The expense-ratio coefficient for the net specification, significant at the 10% level in the main sample, is no longer significant in either alternative. It remains large and significant gross of fees, where the fee added back to returns dominates. The collinearity diagnostics hold as well: the style indicator is nearly perfectly uncorrelated to the four controls (**Table 10** and **Table 11**), and every VIF falls between roughly 1.01 and 1.55 (**Table 12** and **Table 13**), the range already seen in the main sample. All the tables for both outputs are reported the appendix.

[31] A low R^2 does not undermine the regression. The model is designed to test whether investment style predicts alpha after controlling for observable fund characteristics, not to explain all variation in alpha across funds. Most of that variation reflects fund-specific factors such as individual portfolio decisions, stock picks, or luck, none of which the observed characteristics can predict. This is the expected result in the mutual fund literature: Fama and French (2010) show that after risk adjustment, fund alpha is largely indistinguishable from noise. The relevant measure of the model's reliability is the precision of the style coefficient, not the share of total variance explained

6. Conclusion

Over the past two decades, asset managers have handed a growing share of portfolio decisions to computerised models. This shift raises a central question: do systematic, model-driven funds deliver better risk-adjusted performance than more traditional funds relying on human judgement? The existing evidence is fragmented across markets, technologies, and sample periods, and points in conflicting directions. This master's thesis aims at answering this question by comparing quantitative and discretionary US-domiciled equity mutual funds over 2000 to 2024 under a single classification rule and one performance-measurement framework, so that any difference in performance could be attributed to management style rather than to the design choices of the study. Two questions guided the analysis: whether quantitative funds outperform discretionary funds on a risk-adjusted basis (RQ1), and where each management style holds a comparative advantage over the other (RQ2).

The answer to **RQ1** is negative. Quantitative funds underperform their discretionary peers on a risk-adjusted basis. After adjustment by a region-matched six-factor model and after controlling for size, age, fees, flows, and investment region, the average quantitative fund earns about 0.57 percentage points less net alpha per year than its discretionary counterpart, and 0.69 percentage points less gross of fees, both significant at the 1% level. The underperformance is a risk-adjusted result rather than a difference in raw returns, which are statistically indistinguishable across the two groups. It survives all the robustness check: re-estimating the model on funds with at least 24 and at least 60 months of return history leaves the sign and the 1% significance unchanged, and the gap reaches its widest point among the funds with the longest records, which rules out short-lived failures as its source. The finding needs one precision. In absolute terms both styles earn net alpha close to zero, so neither group produces meaningful abnormal returns, in line with Fama and French (2010). The gap just give a winner between two styles that both, on average, fail to beat their risk benchmark once costs are deducted.

Three features of systematic management, each documented in this work, could help explaining why quantitative funds post lower alpha than their discretionary peers. The deficit is visible gross of fees, which places its origin in pre-fee management decisions rather than in costs. Part of the explanation is that nominally active quantitative funds are often less active than their label implies: closet indexing is more common among them and their active share is lower, so investors pay for active management while holding something close to the benchmark (Miguel & Chen, 2025). A second part is crowding. Because many systematic funds scan the same universe on the same published signals and rely on the same data vendors, their positions overlap, and the resulting competition erodes the premia those signals were meant to capture, an effect that intensifies as systematic capital grows (Hu et al., 2024). A third part is adaptability. Rule-based models extrapolate from historical patterns and lose ground

when the informational environment changes, whereas a human analyst can reinterpret a novel situation for which no precedent exists (Birru et al., 2024; Dyer et al., 2025). Together, these mechanisms describe underperformance that is steady and small before fees rather than rare and severe. That pattern matches the narrow, thin-tailed distribution of alphas observed across the group

RQ2 asked where each approach holds an edge, and the results point to clear distinct comparative advantages. The quantitative advantage is twofold: a lower expense ratio, 0.98% against 1.09% per year, and greater consistency, visible in lower alpha dispersion across funds. This consistency has value of its own: for an investor who prefers predictable, benchmark-like performance to the lottery of an individual manager, the systematic profile is attractive even in the absence of alpha. It is worth noting that the less dispersed quantitative alpha could also emerge from their fewer number in the sample. Discretionary funds claim the upper tail: they have more outperformers and earn higher alpha both before and after fees. This is consistent with the literature in which managerial skill, where it exists at all, concentrates among those who exploit soft or specialised information and deep sector knowledge (Birru et al., 2024; Jordan et al., 2022). Quantitative funds are also smaller on average, and size predicts higher alpha in this sample, so part of the gap between the two groups possibly reflects a size disadvantage rather than the investment approach itself. The controls and region fixed effects absorb a little part of the style gap, which holds within size, age, fee, and region. The cost saving of quantitative funds is real, but too small to close their pre-fee deficit.

For investors the implication is discouraging. The evidence gives no support to paying a premium for a quantitative label in the expectation of higher net returns; on this sample the systematic label is associated with lower fees and steadier risk-adjusted performance, not with outperformance. Fund flows are indistinguishable across the two styles, which suggests that investor demand has not yet discriminated on the performance difference the risk-adjusted tests reveal. This has to be paired with the difficulty investors face in assessing true fund quality. Investors seeking outperformance still need to find the right discretionary manager. Systematic funds offer something different: lower fees and steadier performance.

Earlier studies disagreed partly because each adopts their own classification method, performance model, sample, making their findings hard to reconcile. The results of this study align with Miguel and Chen (2021, 2025), whose US-focused evidence pointed toward quantitative underperformance, and extend that finding to an international sample. They also challenge the more positive picture reported for AI-powered funds by Chen and Ren (2022). The results suggest each style fills a different role: discretionary funds retain the edge in generating alpha, quantitative funds win on cost and consistency,

a split consistent with the view that the two approaches are complements rather than rivals (Cao et al., 2024).

These conclusions carry the limitations set out in the methodology, and two of them dictate on how the result should be interpreted. The style classification is static: a fund keeps the label recorded at the end of the sample throughout its history, so a fund that adopted systematic tools midway is treated as quantitative from the start. Because the boundary between the two styles blurred over the period, this likely compresses the measured gap rather than inflating it, which makes the estimated underperformance a conservative figure rather than a product of the design. The assignment of investment regions rests on each fund's stated geographical focus, so for a fund that act outside that region the matched factors mismeasure its risk, and part of what appears as alpha may be uncompensated out-of-region exposure. The study also measures performance at the primary share-class level, so the figures describe that class rather than the full umbrella fund, and a class with a different fee load could show a different net alpha. None of these reverses the direction of the result, but each condition how the results should be interpreted.

Three research directions are motivated by the findings. First, this study treats all quantitative funds as a single group, but the systematic universe is heterogeneous: factor-based strategies, machine learning models, and AI-powered approaches share the quantitative label while operating very differently. Evidence from a narrower sample of AI-powered funds suggests their performance profile differs from the broader quantitative group (Chen & Ren, 2022), which raises the question of whether the underperformance documented here concentrates in specific quantitative strategies or is valid across all of them. Decomposing the quantitative universe by investment technology and process would let future work identify which systematic approaches account for the gap and which, if any, match or exceed discretionary performance. Second, the analysis measures time-weighted fund alpha, which captures fund performance regardless of investor timing. Dollar-weighted returns weight each period by capital under management and reflect what investors actually earn; the two metrics can diverge when flows respond to past performance. The question of whether quantitative and discretionary investors earn different dollar-weighted returns — given that fund flows are indistinguishable across styles in this sample — remains open and carries direct implications for investor welfare. Third, the study documents a cross-sectional performance gap but does not ask whether the same funds drive it across time. Testing performance persistence separately by style and analysing whether the discretionary outperformers persist or represent luck, would extend the analysis that Miguel and Chen (2021) begin on an earlier, narrower sample.

7. Appendix

7.1. Robustness check

The tables below re-run the computation from Sections 6.3 and 6.4 on the two alternative samples discussed in Section 6.5: Robustness A lowers the return-history requirement to 24 months; Robustness B raises it to 60 months. Table specifications match their main-sample counterparts (Tables 5 to 7).

Table 8: Cross-sectional regressions, 24-month sample (Robustness A)

	Dependent variable: FF6 alpha (% p.a.)			
	Net of fees	**Net of fees**	**Net of fees**	**Gross of fees**
	(1)	**(2)**	**(3)**	**(4)**
Quantitative (=1)	**−0.717*** **	**−0.609*** **	**−0.606*** **	**−0.647*** **
	(0.153)	(0.152)	(0.152)	(0.150)
Log TNA		0.253***	0.244***	0.263***
		(0.030)	(0.030)	(0.033)
Age (years)		−0.007	−0.008*	−0.006
		(0.005)	(0.004)	(0.005)
TER (%)		0.246	0.256	1.278***
		(0.182)	(0.186)	(0.174)
Net flow (%)		0.174***	0.176***	0.247***
		(0.044)	(0.044)	(0.056)
Controls	No	Yes	Yes	Yes
Region FE	No	No	Yes	Yes
N	3,121	3,121	3,121	3,040
Adj. R^2	0.004	0.044	0.056	0.074

*Notes: Heteroskedasticity-robust standard errors (HC1) in parentheses. *** p<0.01, ** p<0.05, * p<0.10. One observation per fund. Dependent variable: annualised FF6 alpha in % per year. Sample: Robustness A (minimum 24 months).*

Table 9: Cross-sectional regressions, 60-month sample (Robustness B)

	Dependent variable: FF6 alpha (% p.a.)			
	Net of fees	**Net of fees**	**Net of fees**	**Gross of fees**
	(1)	**(2)**	**(3)**	**(4)**
Quantitative (=1)	**−0.770*** **	**−0.695*** **	**−0.691*** **	**−0.780*** **
	(0.137)	(0.134)	(0.131)	(0.128)
Log TNA		0.228***	0.225***	0.231***
		(0.025)	(0.025)	(0.026)
Age (years)		−0.006	−0.008*	−0.007*
		(0.004)	(0.004)	(0.004)
TER (%)		0.010	0.033	1.112***
		(0.143)	(0.146)	(0.148)
Net flow (%)		0.247***	0.258***	0.265***
		(0.041)	(0.041)	(0.044)
Controls	No	Yes	Yes	Yes
Region FE	No	No	Yes	Yes
N	2,847	2,847	2,847	2,750
Adj. R^2	0.007	0.056	0.072	0.075

*Notes: Heteroskedasticity-robust standard errors (HC1) in parentheses. *** p<0.01, ** p<0.05, * p<0.10. One observation per fund. Dependent variable: annualised FF6 alpha in % per year. Sample: Robustness B (minimum 60 months).*

Table 10:Correlation matrix of the regressors, 24-month sample (Robustness A)

Variable	**Alpha, net (%)**	**Alpha, gross (%)**	**Quant (=1)**	**Log TNA**	**Age (years)**	**TER (%)**
Alpha, net (%)	1					
Alpha, gross (%)	0.927	1				
Quant (=1)	−0.069	−0.081	1			
Log TNA	0.158	0.122	−0.036	1		
Age (years)	0.020	0.024	−0.034	0.478	1	
TER (%)	−0.030	0.105	−0.070	−0.237	0.073	1
Net flow (%)	0.079	0.075	−0.010	−0.258	−0.446	−0.149

Notes: Lower-triangular Pearson correlation matrix. N = 3,040 funds. Sample: Robustness A (minimum 24 months).

Table 11:Correlation matrix of the regressors, 60-month sample (Robustness B)

Variable	Alpha, net (%)	Alpha, gross (%)	Quant (=1)	Log TNA	Age (years)	TER (%)
Alpha, net (%)	1					
Alpha, gross (%)	0.960	1				
Quant (=1)	−0.088	−0.105	1			
Log TNA	0.153	0.091	−0.058	1		
Age (years)	−0.012	−0.006	−0.044	0.439	1	
TER (%)	−0.083	0.092	−0.072	−0.315	0.043	1
Net flow (%)	0.120	0.088	0.013	−0.263	−0.481	−0.160

Notes: Lower-triangular Pearson correlation matrix. N = 2,750 funds. Sample: Robustness B (minimum 60 months).

Table 12: Variance inflation factors, 24-month sample (Robustness A)

Regressor	VIF
Quantitative (=1)	1.010
Log TNA	1.453
Age (years)	1.549
TER (%)	1.148
Net flow (%)	1.286

Notes: VIF from the OLS of net alpha on all regressors without region fixed effects. Sample: Robustness A (minimum 24 months).

Table 13:Variance inflation factors, 60-month sample (Robustness B)

Regressor	VIF
Quantitative (=1)	1.014
Log TNA	1.473
Age (years)	1.528
TER (%)	1.221
Net flow (%)	1.362

Notes: VIF from the OLS of net alpha on all regressors without region fixed effects. Sample: Robustness B (minimum 60 months).

8. References and Bibliography

Abis, S. (2020). *Man vs. Machine: Quantitative and Discretionary Equity Management* (SSRN Scholarly Paper No. 3717371). https://doi.org/10.2139/ssrn.3717371

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and Momentum Everywhere. *The Journal of Finance*, *68*(3), 929–985. https://doi.org/10.1111/jofi.12021

Bain, J. S. (1959). *Industrial Organization*. New York: John Wiley & Sons.

Berk, J. B., & Green, R. C. (2004). Mutual Fund Flows and Performance in Rational Markets. *Journal of Political Economy*, *112*(6), 1269–1295. https://doi.org/10.1086/424739

Birru, J., Gokkaya, S., Liu, X., & Markov, S. (2024). Quants and market anomalies. *Journal of Accounting and Economics*, *78*(1), 101688. https://doi.org/10.1016/j.jacceco.2024.101688

Black, F., & Litterman, R. (1992). Global Portfolio Optimization. *Financial Analysts Journal*, *48*(5), 28–43. https://doi.org/10.2469/faj.v48.n5.28

Cao, S., Jiang, W., Wang, J., & Yang, B. (2024). From Man vs. Machine to Man + Machine: The art and AI of stock analyses. *Journal of Financial Economics*, *160*, 103910. https://doi.org/10.1016/j.jfineco.2024.103910

Capocci, D., & Hübner, G. (2004). Analysis of hedge fund performance. *Journal of Empirical Finance*, *11*(1), 55–89. https://doi.org/10.1016/j.jempfin.2002.12.002

Carhart, M. M. (1997). On Persistence in Mutual Fund Performance. *The Journal of Finance*, *52*(1), 57–82. https://doi.org/10.2307/2329556

Chen, R., & Ren, J. (2022). Do AI-powered mutual funds perform better? *Finance Research Letters*, *47*, 102616. https://doi.org/10.1016/j.frl.2021.102616

Chincarini, L. (2014). The Impact of Quantitative Methods on Hedge Fund Performance. *European Financial Management*, *20*(5), 857–890. https://doi.org/10.1111/eufm.12035

Cogneau, P., & Hübner, G. (2020). International Mutual Funds Performance and Persistence across the Universe of Performance Measures. *Finance*, *41*(1), 97–176. https://doi.org/10.3917/fina.411.0097

Coleman, L. (2023). Explaining mutual fund behavior through the structure-conduct-performance lens. *International Journal of Finance & Economics*, *28*(3), 2874–2884. https://doi.org/10.1002/ijfe.2568

Creamer, G., & Freund, Y. (2010). Automated trading with boosting and expert weighting. *Quantitative Finance*, *10*(4), 401–420. https://doi.org/10.1080/14697680903104113

DeMiguel, V., Gil-Bazo, J., Nogales, F. J., & Santos, A. A. P. (2023). Machine learning and fund characteristics help to select mutual funds with positive alpha. *Journal of Financial Economics*, *150*(3), 103737. https://doi.org/10.1016/j.jfineco.2023.103737

Evans, R. B. (2010). Mutual Fund Incubation. *The Journal of Finance*, *65*(4), 1581–1611.

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

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

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Ferreira, M. A., Keswani, A., Miguel, A. F., & Ramos, S. B. (2013). The Determinants of Mutual Fund Performance: A Cross-Country Study. *Review of Finance*, *17*(2), 483–525. https://doi.org/10.1093/rof/rfs013

Ferson, W. E., & Schadt, R. W. (1996). Measuring Fund Strategy and Performance in Changing Economic Conditions. *The Journal of Finance*, *51*(2), 425–461. https://doi.org/https://doi.org/10.2307/2329367

Giudici, P., & Raffinetti, E. (2023). SAFE Artificial Intelligence in finance. *Finance Research Letters*, *56*, 104088. https://doi.org/10.1016/j.frl.2023.104088

Grinblatt, M., & Titman, S. (1994). A Study of Monthly Mutual Fund Returns and Performance Evaluation Techniques. *The Journal of Financial and Quantitative Analysis*, *29*(3), 419. https://doi.org/10.2307/2331338

Grobys, K., Kolari, J. W., & Niang, J. (2022). Man versus machine: On artificial intelligence and hedge funds performance. *Applied Economics*, *54*(40), 4632–4646. https://doi.org/10.1080/00036846.2022.2032585

Grossman, S. J., & Stiglitz, J. E. (1980). On the Impossibility of Informationally Efficient Markets. *The American Economic Review*, *70*(3), 393–408.

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical Asset Pricing via Machine Learning. *The Review of Financial Studies*, *33*(5), 2223–2273. https://doi.org/10.1093/rfs/hhaa009

Habbal, A., Ali, M. K., & Abuzaraida, M. A. (2024). Artificial Intelligence Trust, Risk and Security Management (AI TRiSM): Frameworks, applications, challenges and future research directions. *Expert Systems with Applications*, *240*, 122442. https://doi.org/10.1016/j.eswa.2023.122442

Harvey, C. R., Liu, Y., & Zhu, H. (2016). … and the Cross-Section of Expected Returns. *Review of Financial Studies*, *29*(1), 5–68. https://doi.org/10.1093/rfs/hhv059

Harvey, C. R., Rattray, S., Sinclair, A., & Hemert, O. V. (2017). Man vs. Machine: Comparing Discretionary and Systematic Hedge Fund Performance. *The Journal of Portfolio Management*, *43*(4), 55–69. https://doi.org/10.3905/jpm.2017.43.4.055

Hu, R., Xiang, W., Zheng, W., & Zhou, K. (2024). Can quantitative investment improve market efficiency?—Evidence from China. *European Financial Management*, *30*(5), 2628–2657. https://doi.org/10.1111/eufm.12485

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance*, *48*(1), 65–91. https://doi.org/https://doi.org/10.2307/2328882

Jensen, M. C. (1968). The Performance of Mutual Funds in the Period 1945-1964. *The Journal of Finance*, *23*(2), 389–416. https://doi.org/https://doi.org/10.2307/2325404

Jordan, B. D., Li, A., & Liu, M. H. (2022). Mutual fund preference for pure-play firms. *Journal of Financial Markets*, *61*, 100719. https://doi.org/10.1016/j.finmar.2022.100719

Kuosmanen, T. (2007). Performance measurement and best-practice benchmarking of mutual funds: Combining stochastic dominance criteria with data envelopment analysis. *Journal of Productivity Analysis*, *28*(1–2), 71–86. https://doi.org/10.1007/s11123-007-0045-7

Mahoney, P. G. (2004). Manager-Investor Conflicts in Mutual Funds. *Journal of Economic Perspectives*, *18*(2), 161–182. https://doi.org/10.1257/0895330041371231

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, *7*(1), 77–91. https://doi.org/10.2307/2975974

Mateus, I. B., Mateus, C., & Todorovic, N. (2019). Review of new trends in the literature on factor models and mutual fund performance. *International Review of Financial Analysis*, *63*, 344–354. https://doi.org/10.1016/j.irfa.2018.12.012

Miguel, A. F., & Chen, Y. (2021). Do machines beat humans? Evidence from mutual fund performance persistence. *International Review of Financial Analysis*, *78*, 101913. https://doi.org/10.1016/j.irfa.2021.101913

Miguel, A. F., & Chen, Y. (2025). How active is your (nominally) actively managed quantitative fund? *International Review of Financial Analysis*, *103*, 104173. https://doi.org/10.1016/j.irfa.2025.104173

Pástor, Ľ., Stambaugh, R. F., & Taylor, L. A. (2015). Scale and skill in active management. *Journal of Financial Economics*, *116*(1), 23–45. https://doi.org/10.1016/j.jfineco.2014.11.008

The stockmarket is now run by computers, algorithms and passive managers. (2019). *The Economist*. https://www.economist.com/briefing/2019/10/05/the-stockmarket-is-now-run-by-computers-algorithms-and-passive-managers?utm_medium=cpc.adword.pd&utm_source=google&ppccampaignID=18151738051&ppcadID=&utm_campaign=a.22brand_pmax&utm_content=conversion.direct-response.anonymous&gclsrc=aw.ds&gad_source=1&gad_campaignid=18151761343&gclid=Cj0KCQjwlLDQBhDjARIsAPlIefGJA9PZ7JkjGi27q4UdkU2Wfa_k-M_2ozhc8P3Fu4kDApnKJTsh-EAaAr-hEALw_wcB

Verma, S., Rao, P., & Kumar, S. (2024). Is investing inherently emotionally arousing process? Fund manager perspective. *Qualitative Research in Financial Markets*, *16*(2), 380–400. https://doi.org/10.1108/QRFM-09-2022-0153

Welch, B. L. (1947). The Generalization of `Student's' Problem when Several Different Population Variances are Involved. *Biometrika*, *34*(1/2), 28. https://doi.org/10.2307/2332510

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, *48*(4), 817. https://doi.org/10.2307/1912934

Yan, X. (Sterling). (2008). Liquidity, Investment Style, and the Relation between Fund Size and Fund Performance. *Journal of Financial and Quantitative Analysis*, *43*(3), 741–767. https://doi.org/10.1017/S0022109000004270

Yin, C. (2016). The Optimal Size of Hedge Funds: Conflict between Investors and Fund Managers. *The Journal of Finance*, *71*(4), 1857–1894. https://doi.org/10.1111/jofi.12413

Zhang, J., & Huang, W. (2021). Option hedging using LSTM-RNN: An empirical analysis. *Quantitative Finance*, *21*(10), 1753–1772. https://doi.org/10.1080/14697688.2021.1905171

Zhang, J., Peng, Z., Zeng, Y., & Yang, H. (2023). Do big data mutual funds outperform? *Journal of International Financial Markets, Institutions and Money*, *88*, 101842. https://doi.org/10.1016/j.intfin.2023.101842

EXECUTIVE SUMMARY

Computerised models have taken over a growing share of investment decisions. This master's thesis asks whether the shift has paid off by comparing US-domiciled actively managed equity mutual funds from 2000 to 2024. The sample covers 3,028 funds, 222 quantitative (model-driven) and 2,806 discretionary (human-managed). Lipper's prospectus-based flag identifies the model-driven funds. A six-factor Fama–French–Carhart model measures performance, with factors matched to each fund's investment region. The first step produces a risk-adjusted return (alpha) for each fund. The second step compares those alphas between the two styles, controlling for size, age, fees, flows and region.

Net monthly returns are almost identical at around 0.77% for both styles. The difference shows up in alpha. After controlling for size, age, fees, flows and investment region, quantitative funds earn on average 0.57 percentage points less net alpha per year than discretionary funds, and 0.69 points less gross of fees. Both gaps are significant at the 1% level. The result holds when the minimum return history is shortened to 24 months or extended to 60 and is the widest among the longest-lived funds.

For an investor expecting a systematic fund to outperform, the data offers no support. Neither style earns high positive alpha on average. The key difference is that quantitative funds offer lower fees and more predictable outcomes, while discretionary funds carry more dispersion but also more outperformers and a higher average alpha, net and gross of fees. The two approaches appear complementary rather than rivals.

MOTS-CLÉS/KEYWORDS: Mutual Fund, quantitative, discretionary, performance, equity, alpha, investment, management

NOMBRE DE MOTS/WORD COUNT: 16400

Ecole de Gestion de l'Université de Liège