Proceedings of ICEMGD 2026 Symposium: Data-Driven Business Management, Corporate Governance, and Organizational Performance
DOI: 10.54254/2754-1169/2026.LD36191

Active Mutual Fund Performance, Benchmark Exposure, and Selection Skill in China's A-Share Market

Minze Li

School of Finance, Capital University of Economics and Business, Beijing, China
32023110153@cueb.edu.cn

Abstract. This paper tests whether active funds can generate extra value beyond passive benchmarks, and how such value should be understood. The paper uses monthly fund returns and A-share indices, including CSI300, CSI500 and CSI800, to examine fund performance from several perspectives, including aggregate fund returns, cross-sectional differences, benchmark regressions, quintile sorting, market-neutral adjustment, and a 2025 out-of-sample test. The results show that the performance of the equal-weighted (EW) active-fund portfolio is better than that of the main benchmarks, but this result masks large differences across individual funds. Many funds have positive alpha estimates, but far fewer funds have statistically significant alpha. A sorting strategy based on past 12-month returns is able to create a clear gap between high- and low-ranked funds, and this gap remains visible in 2025. Nevertheless, after market-neutral and dual-benchmark adjustments, part of the top-quintile return is absorbed by market and style exposure. Overall, active funds appear to add value, but the evidence points to a mixture of exposure, persistence, and smaller residual alpha rather than pure manager skill.

Keywords: active mutual funds, benchmark exposure, performance persistence, fund selection, A-share market

1. Introduction

Whether active mutual funds are able to create value through managers' skill relative to passive benchmarks remains an important question in empirical asset management research. Active fund managers may add value by acquiring and processing information, selecting mispriced securities, and adjusting portfolios across shifting market conditions. Simultaneously, apparent outperformance is difficult to interpret. Higher returns may reflect greater risk exposure rather than true skill, and positive alpha estimates may come from benchmark mismatch, statistical noise and luck, not from the persistent skill performance [1-3]. Therefore, active management should not be assessed simply by comparing raw returns with benchmarks.

This problem becomes more complex in China's A-share market, because of its distinctive features, including large market scale, high volatility, substantial portion of individual investors, and unique political environment, it is more important to study the value creation ability of active funds as an important group of institutional investors. At the same time, it's hard to select a benchmark, because there is no unified index that captures universe of active A-share funds, the exposure of

DOI: 10.54254/2754-1169/2026.LD36191

active funds to large-cap and mid-cap indices will be very different. Relying on one single benchmark may lead to biased estimates of fund returns due to different style tilts. Therefore, analyzing the returns of active funds requires a rigorous analysis of benchmark exposure, cross-sectional selection, residual alpha [4-6].

First, the analysis begins by comparing the equal-weighted active fund portfolio with a passive benchmark, and then examine the cross-section of the active fund industry to assess the selection risk of investors. Based on this, single-benchmark regression and dual-benchmark regression are used to separate benchmark-related exposure from residual alpha. Finally, this paper evaluates the persistence of fund performance by testing whether past performance predicts future performance, which helps assess whether funds have genuine skill.

This paper contributes to the empirical analysis of active fund performance in these ways. First, this paper shows that the outperformance of equal-weighted active-fund portfolio relative to passive benchmarks should be interpreted with caution, and that much of it comes from benchmark and style exposure, diversification, and possibly some skill. Second, this paper distinguishes positive alpha estimates from statistically significant alpha, and how to explain why most of the alpha is statistically insignificant is a more critical issue. Third, based on a series of sorting signal tests, this paper shows that past performance can help identify stronger funds, and the performance of funds has a degree of persistence.

2. Data and empirical design

In the A-share market, the CSI300 is usually used as a benchmark for large-cap stocks, the CSI500 represents more mid-cap exposure, and the CSI800 includes certain small-cap stocks on this basis. Unlike CSI800, the dual-benchmark regression estimates separate large-cap and mid-cap loadings. If the CSI300 is used to evaluate a fund that tilts toward mid-cap stocks, the additional return may be attributed to alpha, resulting in the overestimation of the fund's alpha. Conversely, if the CSI800 is selected as the benchmark for the same fund, it is likely to absorb part of the fund's intercept, which understates the performance.

Therefore, this paper will use both single-benchmark and dual-benchmark regressions in the following research. The single-benchmark regression first provides simple alpha results, while the inclusion of the two-benchmark regression reduces the problem of benchmark mismatch. Since most funds are tilted towards mid- and large-cap stocks, using the dual regressions with CSI300 and CSI500 is sufficient to reduce most of the mismatch problems.

2.1. Data and sample construction

The analysis uses two primary datasets, which are monthly index returns for major A-share benchmarks and monthly returns of active funds. The benchmarks include CSI300 (large-cap), CSI500 (mid-cap), and CSI800 (broad market). Fund returns are recorded at the individual fund level, and zeros are treated as missing observations.

To compare funds and indices, all return series are transformed into a consistent monthly time index. For funds, the panel is unbalanced: funds enter and exit over time, so the effective number of funds varies by month. All aggregation and regression steps use only the months in which the relevant series are simultaneously observed.

DOI: 10.54254/2754-1169/2026.LD36191

2.2. Equal-weighted industry portfolio and performance metrics

To characterize active fund industry-level performance, construct an equal-weighted active-fund portfolio. In each month, take the simple average return across funds with non-missing returns in that month.

For each return series, include benchmarks and EW fund portfolio, compute a common set of performance metrics: annualized arithmetic mean return (Ann mean), annualized volatility (Ann vol), Sharpe ratio (assuming the risk-free rate rf=0), maximum drawdown (Max DD), Calmar ratio, win rate, and tail risk measures such as VaR95 estimated from the empirical monthly return distribution.

2.3. Benchmark attribution models

To separate benchmark exposure from residual performance, estimate two classes of regression:

Single-benchmark regression:

$$r_i = \alpha_i + \beta_i r_b + \varepsilon_i \tag{1}$$

where r_b is CSI300, CSI500, or CSI800.

Dual-benchmark regression:

$$r_i = \alpha_i + \beta_{i,300} r_{300} + \beta_{i,500} r_{500} + \varepsilon_i \tag{2}$$

This regression aims to absorb style risks, especially when a fund tilts toward mid-cap stocks, thereby reducing benchmark mismatch that would inflate alpha under a single large-cap benchmark.

Because a risk-free series is not included, all regressions use raw returns; therefore, alpha should be interpreted as a benchmark-relative intercept, not a strict excess-return capital asset pricing model (CAPM) alpha. For the fund-level cross-sectional regression summaries reported in Section 3, funds are required to have at least 24 overlapping monthly observations.

3. Aggregate performance and benchmark attribution

3.1. Aggregate active-fund performance versus benchmarks

Numerically, the EW active-fund portfolio delivers an annualized arithmetic mean return of 14.57%, versus 9.81% (CSI300), 12.48% (CSI500), and 10.14% (CSI800). Risk-adjusted performance is also higher, EW Sharpe is 0.608 compared with 0.357–0.405 for the benchmarks, while EW Max DD is −56.9% versus −69% to −70.8% for the benchmarks. Tail risk measured by VaR95 is broadly comparable (EW 10.12% vs benchmarks around 10.23%–13.18%), while EW win rate is higher (0.589 vs 0.536–0.549) (Table 1).

Table 1. Aggregate and cross-sectional fund performance

Panel A: Aggregate performance of the EW fund portfolio and benchmarks

Name	Ann mean	Ann vol	Max DD	Sharpe	Calmar	Win rate	VaR95	Skew	Ex. Kurt.
EW fund	0.1457	0.2396	-0.5692	0.6082	0.2169	0.5890	0.1012	-0.1635	1.1422

DOI: 10.54254/2754-1169/2026.LD36191

Panel A. (continued)

CSI300	0.0981	0.2753	-0.7075	0.3565	0.0876	0.5485	0.1023	0.0418	1.4620
CSI500	0.1248	0.3083	-0.6927	0.4049	0.1158	0.5359	0.1318	0.0607	1.3422
CSI800	0.1014	0.2754	-0.7005	0.3683	0.0933	0.5443	0.1027	-0.0362	1.3503

Panel B: Cross-sectional distribution of individual-fund performance

Metrics	p05	p25	Median	p75	p95	Mean	SD
Ann mean	-0.0164	0.0439	0.0800	0.1132	0.1577	0.0763	0.0638
Ann vol	0.1648	0.2008	0.2340	0.2624	0.3040	0.2331	0.0468
Sharpe	-0.0785	0.1906	0.3419	0.4853	0.6962	0.3256	0.3605
Max DD	-0.6485	-0.5662	-0.4823	-0.3831	-0.2427	-0.4679	0.1298
Win rate	0.4604	0.5060	0.5339	0.5608	0.6058	0.5336	0.0502
VaR95	0.0657	0.0835	0.0988	0.1144	0.1364	0.0995	0.0228
Skew	-0.4340	-0.1166	0.0972	0.3274	0.7390	0.1187	0.3772
Ex. Kurt.	-0.3957	0.2455	0.9248	1.7669	3.1050	1.1071	1.2574

This overall outperformance of the equal-weighted active fund portfolio should not be interpreted as evidence that the average active manager has better stocks-selection ability. First, the equal-weighted portfolio as a whole cannot be matched by a single benchmark, and the apparent excess returns are likely to reflect style exposure rather than skill. Second, an equal-weighted portfolio of all funds diversifies the erroneous selection of a particular fund, so the portfolio may be stronger than most individual funds. Third, even if the management component exists, it is difficult to identify how much comes from the first two aspects and how much comes from skill. This is consistent with the equilibrium-accounting view, which argues that such excess performance needs to be interpreted with caution [3, 6, 7].

3.2. Cross-sectional fund performance and selection risk

Within the active-fund universe, cross-sectional dispersion is substantial. The median fund has annually mean of 8.00% and mean of 7.63%, but the distribution is wide, the 5th percentile is −1.64% and the 95th percentile reaches 15.77%. Risk-adjusted performance shows the same pattern, Sharpe has a median of 0.342 (mean 0.326), with a 5th–95th percentile range of −0.079 to 0.696. Drawdowns are also dispersed (median Max DD −48.2%, 5th percentile −64.9%, 95th percentile −24.3%), consistent with meaningful selection risk. Moreover, excess kurtosis is elevated (median 0.925, 95th percentile 3.105), indicating heavier tails than a normal benchmark. (Panel B of Table 1).

The dispersion implies that active vs passive is not a uniform statement. Even if the EW industry portfolio looks strong, investors selecting individual funds face selection risk, which means outcomes depend heavily on where a chosen fund lies in the cross-sectional distribution.

DOI: 10.54254/2754-1169/2026.LD36191

3.3. Benchmark attribution: single- and dual-benchmark evidence

3.3.1. Single-benchmark alpha & beta attribution

At the aggregate level, single-benchmark regressions show substantial benchmark exposure, with betas of 0.689–0.797. The EW portfolio also has annualized alphas of 5.84%–6.82%, t-statistics of 2.33–2.92, and R^2 values of 0.788–0.842, and all three alpha estimates are statistically significant (Table 2).

Table 2. Benchmark regression results and fund-level cross-sectional

Panel A: Aggregate portfolio regressions

	CSI300	CSI500	CSI800	CSI300+CSI500
Alpha	0.0682*** (2.73)	0.0584** (2.33)	0.0631*** (2.92)	0.0567*** (2.69)
Beta on CSI300	0.7721*** (29.59)			0.4186*** (9.99)
Beta on CSI500		0.6886*** (29.49)		0.3703*** (9.90)
Beta on CSI800			0.7972*** (35.37)	
R-squared (Adjusted R-squared)	0.7891	0.7880	0.8424	0.8516 (0.8503)

Panel B: Single-benchmark fund-level cross-section

Benchmark	Median(α)	p25(α)	p75(α)	α>0 ratio	α sig ratio (5%)	Median(β)	p25(β)	p75(β)
CSI300	0.0600	0.0307	0.0873	0.9036	0.1757	0.8504	0.7752	0.9394
CSI500	0.0655	0.0370	0.0967	0.9187	0.2447	0.8263	0.7233	0.9264
CSI800	0.0616	0.0332	0.0900	0.9091	0.2235	0.9157	0.8378	1.0090

Panel C: Dual-benchmark fund-level cross-sectional summary

Parameter	Median	p25	p75	α>0 ratio	α sig ratio (5%)
Alpha	0.0614	0.0348	0.0922	0.9202	0.2753
Beta on CSI300	0.2971	0.0528	0.5351	—	—
Beta on CSI500	0.6148	0.3448	0.8780	—	—

3.3.2. Cross-sectional analysis

At the individual-fund level, the median annualized alpha is 6.00% under CSI300, 6.55% under CSI500, and 6.16% under CSI800, with interquartile ranges of 3.07%–8.73%, 3.70%–9.67%, and

DOI: 10.54254/2754-1169/2026.LD36191

3.32%–9.00%, respectively. Mean alphas are similar in magnitude (about 5.92%–6.60% depending on benchmark), but the tails are wide: the 5th percentile is around −1.7% to −2.2%, while the 95th percentile is about 13.7%–14.5%. Betas are also dispersed (median β ≈ 0.83–0.92, with a 5th–95th percentile band roughly 0.59–1.09 depending on benchmark), confirming that funds differ materially in market exposure. Importantly, "alpha is positive" is far more common than "alpha is statistically strong": roughly 90%–92% of funds have α>0, but only about 17.6% (CSI300) to 24.5% (CSI500) have t(α)>1.96.

The fact that around 90% of funds have alpha above zero, while only a much smaller fraction produce statistically significant alpha, suggests that point estimates alone overstate the strength of active management. In this sense, the results are consistent with the broader literature that emphasizes the need to distinguish skill from luck in mutual fund performance evaluation [2, 3].

3.3.3. Two-benchmark regression (CSI300 + CSI500)

CSI300 and CSI500 are highly correlated (0.8521), which may inflate the standard errors of beta estimates. Nevertheless, including CSI500 remains economically useful because it reduces benchmark mismatch for funds that tilt away from large-cap stocks.

The two-benchmark regression provides a cleaner style decomposition for the EW industry portfolio. Specifically, the EW portfolio loads on both large-cap and mid-cap benchmarks, with $\beta_{\text{CSI300}} = 0.4186$ and $\beta_{\text{CSI500}} = 0.3703$ (sum ≈ 0.789) and delivers an annualized intercept of $\alpha = 5.67\%$ with $t(\alpha) = 2.69$. Model fit is strong with $R^2 = 0.852$, indicating that most of the EW portfolio's variation is explained by these two benchmarks (Table 2).

Cross-sectionally, Table 2 confirms a pronounced mid-cap tilt in the active-fund universe. The median fund has β_{CSI500} =0.615 (interquartile range of 0.345–0.878) and β_{CSI300} =0.297 (interquartile range of 0.053–0.535), indicating that mid-cap exposure is typically larger than large-cap exposure once style is decomposed. Alpha point estimates remain mostly positive: the median annualized alpha is 6.14% (interquartile range of 3.48%–9.22%) and about 92.0% of funds have α>0; however, only 27.5% meet a rough statistically significant which means t(α)>1.96.

The two-benchmark results suggest that part of the alpha measured under a single benchmark is likely to reflect benchmark mismatch rather than pure abnormal performance. Once both large-cap and mid-cap benchmark exposures are included, the regression provides a cleaner decomposition of style risk, and the remaining intercept becomes a more demanding measure of residual performance.

4. Fund sorting and 2025 out-of-sample validation

4.1. Dynamic tranche portfolios based on trailing performance

This section evaluates whether a simple, implementable fund-selection rule can generate economically meaningful performance differences when applied to the cross-section of active A-share mutual funds. Each month, funds are re-sorted dynamically using a performance signal defined as the average monthly return over the past 12 months [8-11]. Funds are then assigned into quintiles (Q1–Q5) and held for one month with equal weights within each tranche, before the process repeats next month.

Importantly, the superior long-run wealth of Q5 is not accompanied by worse downside experience. Panel A of Table 3 reports that annualized mean returns increase from 12.77% (Q1) to 18.51% (Q5), while the Sharpe ratio rises from 0.53 to 0.70, meanwhile, maximum drawdown improves from approximately −44.6% (Q1) to −41.5% (Q5), compared with −55.9% for CSI300.

DOI: 10.54254/2754-1169/2026.LD36191

Table 3. Quintile portfolio performance: in-sample and 2025 out-of-sample results

Panel A: In-sample dynamic quintile portfolio performance

Portfolio	Ann. mean (%)	Ann. vol. (%)	Sharpe	Max. DD (%)	Win rate (%)
Q1	12.7673	24.1877	0.5278	-44.6208	55.7047
Q2	14.9850	24.8280	0.6036	-45.1372	59.7315
Q3	14.9517	24.9646	0.5989	-44.6517	59.0604
Q4	16.7684	25.5081	0.6574	-43.1411	60.4027
Q5	18.5099	26.3959	0.7012	-41.5112	63.0872
CSI300	12.0604	27.8519	0.4330	-55.8669	56.3758

Panel B: 2025 out-of-sample dynamic quintile portfolio performance

Portfolio	Ann. return	Ann. vol	Sharpe	Max. DD	Win rate	Skew	Ex. Kurt.
CSI300	0.1651	0.1328	1.2434	-0.0307	0.5455	0.8397	0.1524
Q1	0.1245	0.0655	1.9005	-0.0191	0.7273	0.4066	-0.8412
Q2	0.2279	0.1293	1.7620	-0.0291	0.6364	1.2486	0.6988
Q3	0.2734	0.1534	1.7830	-0.0455	0.6364	0.7589	-0.4731
Q4	0.3344	0.1814	1.8435	-0.0449	0.7273	0.8467	-0.2138
Q5	0.3617	0.1989	1.8181	-0.0464	0.5455	0.8993	-0.2466

To isolate the "top-minus-bottom" spread implied by the sorting rule, a series Q5-Q1 can be constructed from monthly tranche returns. This spread delivers an annualized mean of approximately 5.74% with Sharpe around 0.74, and a comparatively moderate maximum drawdown near −11.3%.

The quintile-sorting results suggest that the strategy is capturing more than one mechanism at the same time. One possible channel is performance persistence: funds that performed well over the previous 12 months continue to perform relatively well in the following month. At the same time, the Q5 portfolio also has higher volatility than lower-ranked portfolios, which implies that the strategy is not isolating pure skill alone. A more balanced reading is that the sorting rule captures a combination of persistence in past performance and exposure to rewarded risks or style tilts. Therefore, the strong Q5 outcome should be interpreted as evidence of useful cross-sectional selection, but not as direct proof that the top quintile represents pure manager alpha.

Since the sorting rule generates economically meaningful spread in sample, the next question is whether this outperformance mainly reflects broad market exposure or whether a smaller benchmark-adjusted component also remains.

4.2. Market-adjusted portfolio performance

To examine how much of the strong performance of the top tranche is driven by broad market exposure, and how much remains after the market component is removed, this section constructs a

DOI: 10.54254/2754-1169/2026.LD36191

market-neutral version of the Q5 portfolio by subtracting benchmark returns from the monthly Q5 returns. The baseline specification uses a β of 1.0.

The results show that the raw Q5 portfolio is indeed strong, but a substantial part of its performance is linked to market-wide movements. The raw Q5 portfolio delivers an annualized mean return of about 18.51%, with annualized volatility of about 26.40%, a Sharpe ratio of about 0.70, and a maximum drawdown of about -41.5%. After subtracting the benchmark with a hedge coefficient of 1.0, the market-neutral version still remains positive, but the magnitude falls noticeably: annualized mean return declines to about 6.45%, annualized volatility falls to about 14.56%, the Sharpe ratio declines to about 0.44, and maximum drawdown improves to about -23.8%.

The adjustment shows that Q5's raw performance is not fully market-independent. Nevertheless, the market-adjusted series retains a positive average return and Sharpe ratio, indicating that Q5 contains both a market-related component and a residual component.

Overall, the market-adjusted exercise shows that the strong raw Q5 result is partly market-driven but not fully exhausted by benchmark exposure. The coefficient of 1.0 provides a conservative full-benchmark subtraction in the in-sample exercise.

4.3. 2025 out-of-sample validation

4.3.1. Dynamic out-of-sample design and results

This section evaluates whether the fund selection rule identified in the in-sample period remains effective in 2025 and keeps the dynamic framework unchanged. The out-of-sample window covers January to November 2025, and CSI300 is used as the benchmark.

According to Figure 1, the dynamic strategy continues to produce a return gradient across quintiles in the out-of-sample period. By the end of November 2025, cumulative net asset value (NAV) reaches about 1.114 for Q1, 1.207 for Q2, 1.248 for Q3, 1.303 for Q4, and 1.327 for Q5, while the CSI300 ends at 1.150. The ordering is broadly monotonic, and the top quintile outperforms other quintiles and the benchmark. This indicates that the 12-month average return signal retains substantial cross-sectional predictive content out of sample.

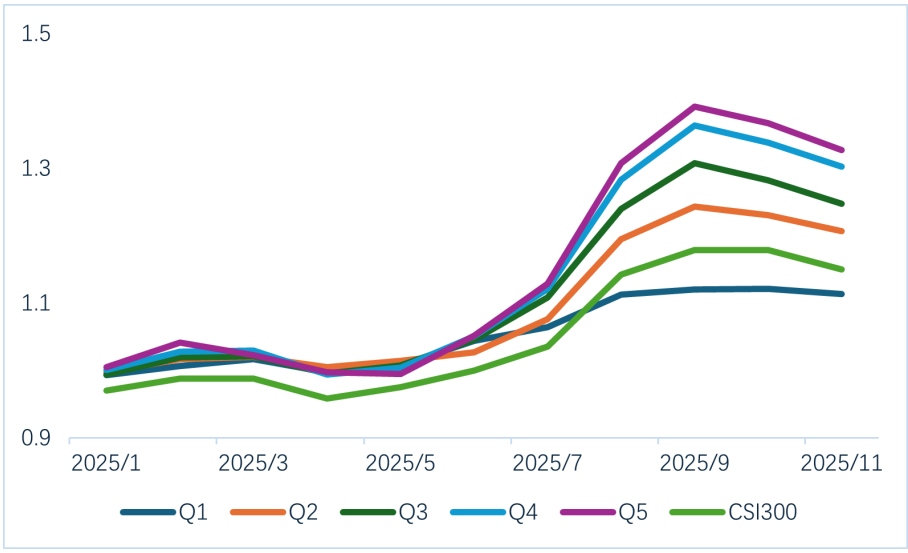

Figure 1. Cumulative NAV of dynamic quintile portfolios and CSI300 in 2025

DOI: 10.54254/2754-1169/2026.LD36191

The same pattern is confirmed by Panel B of Table 3. Q5 delivers an annualized return of 36.17%, compared with 12.45% for Q1 and 16.51% for CSI300.

However, the outperformance is stronger in raw return terms than in risk-adjusted terms. Q5 has the highest annualized return, but its annualized volatility also rises to 19.89%, above both Q1 and the benchmark. Thus, the Sharpe ratio of Q5 is 1.82, which is strong but not the highest. Q1 records a Sharpe ratio of 1.90, and Q4 reaches 1.84. This suggests that the signal is more effective at sorting high-return funds than at identifying the most efficient portfolios on a risk-adjusted basis.

4.3.2. Extension

In addition to this dynamic specification, this section also considers an extension, which funds are first restricted to those with at least 24 months of available return history by December 2024. They are then ranked by the arithmetic mean of all available historical monthly returns up to December 2024, grouped into quintiles, and held through 2025 without rebalancing. This extension is designed to examine whether long-history winner funds continue to outperform out of sample.

Q5 records an annualized return of 33.58%, compared with 18.82% for Q1 and 16.51% for the benchmark.

4.3.3. Benchmark-adjusted performance

The dynamic top quintile remains leading after a benchmark adjustment. After subtracting 0.8 times the CSI300 return, the adjusted Q5 still earns an annualized return of 21.36%, while volatility falls to 10.86%. As a result, the Sharpe ratio improves to 1.97.

However, a single-benchmark adjustment may still understate broader style exposure. For that reason, the next section extends the benchmark adjustment to a dual-benchmark framework using CSI300 and CSI500 jointly.

4.3.4. Interpretation

The out-of-sample evidence confirms that the past 12-month return signal retains predictive power in 2025, but the message should remain balanced. On the one hand, the monotonic increase in cumulative NAV and annualized return from Q1 to Q5 shows that the signal continues to sort funds in a meaningful way out of sample. On the other hand, the top quintile also carries higher volatility, and the Sharpe ratio is not strictly increasing across quintiles. This means that the signal works more clearly as a raw-return sorting device than as a fully risk-adjusted ranking rule. In economic terms, the out-of-sample results support persistence in past performance, but they also suggest that part of Q5's strength remains linked to risk-taking rather than to a pure alpha component. This interpretation is broadly consistent with the prior evidence on performance persistence in the Chinese A-share mutual fund market, while also being more cautious about the role of risk [12-14].

4.4. Dual-benchmark-adjusted test

While the previous subsection shows that the top quintile retains meaningful performance after a single-benchmark adjustment, such evidence may still partly reflect exposure to broader market style factors. To address this concern, the analysis is extended to a dual-benchmark setting based on CSI300 and CSI500. More specifically, the pre-2025 historical return series of the top-quintile portfolio is regressed on the two benchmark returns to estimate its benchmark loadings, and these estimated coefficients are then fixed and carried into the 2025 out-of-sample window. The resulting

DOI: 10.54254/2754-1169/2026.LD36191

adjusted series therefore removes both large-cap and mid-cap benchmark components from the raw Q5 return.

The raw Q5 portfolio delivers an annualized return of 36.17% in the 2025 out-of-sample period, with an annualized volatility of 19.89% and a Sharpe ratio of 1.82. After adjusting only for CSI300, the annualized return declines to 21.36%, while volatility falls to 10.86%, producing a slightly higher Sharpe ratio of 1.97. However, once the benchmark structure is extended to CSI300 and CSI500 jointly, the annualized return falls further to 16.59%, which is almost identical to the benchmark return of 16.51%. This suggests that a substantial part of the raw outperformance can be absorbed once both large-cap and mid-cap benchmark exposures are taken into account.

Nevertheless, the dual-benchmark-adjusted series remains slightly stronger in risk-adjusted terms than the market benchmark. Its annualized volatility declines to 8.33%, compared with 13.28% for CSI300, while its Sharpe ratio rises to 1.99, well above the benchmark Sharpe ratio of 1.24. In addition, the maximum drawdown narrows to -2.52%, compared with -3.07% for CSI300. Therefore, although the dual-benchmark adjustment largely removes the raw excess return of Q5, it does not eliminate the possibility that the top quintile still retains a modestly superior risk-return profile.

4.5. Alternative performance signals for fund sorting

To examine whether out-of-sample fund selection depends on how past performance is measured, this section compares three alternative sorting signals while keeping the method of the portfolio construction procedure unchanged [3, 11, 15]. In all cases, funds are sorted monthly, the lookback window is fixed at 12 months, portfolios are formed into five quintiles, and the next month's returns are equally weighted within each quintile. The out-of-sample evaluation window covers January to November 2025. The first signal is the Sharpe ratio, which captures total risk-adjusted past performance. The second signal is alpha estimated from a single-benchmark regression using CSI300. The third signal is alpha estimated from a dual-benchmark regression using both CSI300 and CSI500.

The strongest result remains the original raw-return sorting from the previous section, whose Q5 portfolio records an annualized return of 36.17% and ends the sample with a cumulative NAV of 1.327. Among the alternative signals examined here, Alpha300 performs best, with a Q5 annualized return of 34.33% and a terminal NAV of 1.311. By comparison, the Sharpe-based Q5 produces an annualized return of 29.67% and a terminal NAV of 1.269, while the Alpha300-500 based Q5 yields 31.58% and a terminal NAV of 1.286. All three exceed CSI300, whose annualized return is 16.51% and terminal NAV is 1.150, but the magnitude of outperformance differs clearly across signals.

The most important difference across sorting signals lies in their ability to generate cross-sectional spread. The annualized Q5-Q1 spread is 8.45% for Alpha300, compared with only 3.09% for the Sharpe signal and 1.51% for Alpha300-500. This indicates that Alpha300 is not only associated with a relatively strong Q5 portfolio, but also produces the clearest separation between winners and losers in the out-of-sample period. By contrast, once both CSI300 and CSI500 are controlled for in the alpha estimation stage, the remaining signal becomes more conservative and less powerful in distinguishing future high- and low-performing funds.

In the original raw-return sorting exercise, the Q5-Q1 spread is about 5.74% per year. By contrast, the Alpha300 signal produces a larger annualized Q5-Q1 spread of 8.45% in the 2025 out-of-sample test. Therefore, if the criterion is the absolute performance of the top portfolio alone, the original raw-return signal remains the strongest; however, if the focus is on cross-sectional sorting

DOI: 10.54254/2754-1169/2026.LD36191

ability, Alpha300 appears more effective because it generates a wider separation between winners and losers.

At the same time, the weaker result for Alpha300-500 is itself informative. Once both large-cap and mid-cap benchmark exposures are controlled for, the alpha signal becomes more stringent, but its cross-sectional discriminatory power also declines.

5. Conclusion

The overall picture that emerges from the results is one of qualified active-fund advantage rather than a simple active versus passive. At the aggregate level, the equal-weighted active-fund portfolio outperforms the main benchmarks in both return and several risk-adjusted dimensions. Yet this should not be read as evidence that manager skill is uniformly strong across the industry. A more convincing interpretation is that the aggregate result reflects a mixture of diversification across funds, benchmark mismatch, style exposure, and some genuine active-management value. Fund-level performance is highly uneven, with substantial dispersion in annualized return, Sharpe ratio, drawdown, and alpha. That dispersion means that a strong aggregate outcome can coexist with considerable selection risk at the individual-fund level. Positive alpha estimates are common under both single-benchmark and dual-benchmark specifications, but statistically reliable alpha is much less widespread. The two-benchmark results sharpen this interpretation further by showing that once both large-cap and mid-cap exposures are accounted for, the remaining residual component is harder to sustain and more limited in scope.

At the same time, the sorting evidence shows that past performance is not irrelevant. The quintile strategy based on trailing returns generates a meaningful spread between top and bottom funds, and the top quintile outperforms both the bottom quintile and the benchmark in raw return terms. Even so, the strategy should not be interpreted as isolating pure alpha. What remains after those adjustments is smaller, but not zero, which implies that the strongest funds combine rewarded exposure with a more limited residual component. The out-of-sample evidence points in the same direction. The 12-month return signal continues to work in 2025, the dynamic specification preserves a clear performance gradient across quintiles, while the fixed specification retains a meaningful top-minus-bottom performance gap. However, the outperformance is clearer in raw return than in fully risk-adjusted terms, since the Sharpe ratio is not strictly increasing across the ranking. Among the proxies examined, Alpha300 appears to deliver the strongest balance between economic interpretation and practical sorting ability, while Sharpe and Alpha300-500 are less effective in separating future winners from losers.

Taken together, the results suggest that active-fund performance in the A-share market is best understood as a layered outcome. One layer comes from broad market and style exposure. Another comes from cross-sectional selection, which is visible in the sorting results. Beyond these, a smaller residual alpha appears to remain, but it is clearly more limited than the raw performance of the top portfolios might initially suggest. The main conclusion, therefore, is not that active management generates large and pervasive skill-based outperformance, but that it creates value through a combination of exposure, persistence, and a narrower residual component that survives stricter benchmark controls.

References

[1] Kosowski, R., Timmermann, A., Wermers, R., & White, H. (2006). Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis. *Journal of Finance, 61*(6), 2551–2595.

DOI: 10.54254/2754-1169/2026.LD36191

[2] Barras, L., Scaillet, O., & Wermers, R. (2010). False discoveries in mutual fund performance: Measuring luck in estimated alphas. *Journal of Finance, 65*(1), 179–216.
[3] Fama, E. F., & French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *Journal of Finance, 65*(5), 1915–1947.
[4] 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.
[5] Daniel, K., Grinblatt, M., Titman, S., & Wermers, R. (1997). Measuring mutual fund performance with characteristic-based benchmarks. *Journal of Finance, 52*(3), 1035–1058.
[6] Wermers, R. (2000). Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses. *Journal of Finance, 55*(4), 1655–1695.
[7] Berk, J. B., & Green, R. C. (2004). Mutual fund flows and performance in rational markets. *Journal of Political Economy, 112*(6), 1269–1295.
[8] Grinblatt, M., & Titman, S. (1992). The persistence of mutual fund performance. *Journal of Finance, 47*(5), 1977–1984.
[9] Hendricks, D., Patel, J., & Zeckhauser, R. (1993). Hot hands in mutual funds: Short-run persistence of relative performance, 1974–1988. *Journal of Finance, 48*(1), 93–130.
[10] Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance, 48*(1), 65–91.
[11] Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*(1), 57–82.
[12] Rao, Z.-u.-R., Iqbal, A., & Tauni, M. Z. (2016). Performance persistence in institutional investment management: The case of Chinese equity funds. *Borsa Istanbul Review, 16*(3), 146–156.
[13] Gao, J., O'Sullivan, N., & Sherman, M. (2017). Performance persistence in Chinese securities investment funds. *Research in International Business and Finance, 42*, 1467–1477.
[14] Rao, Z.-u.-R., Ahsan, T., Tauni, M. Z., & Umar, M. (2018). Performance and persistence in performance of actively managed Chinese equity funds. *Journal of Quantitative Economics, 16*(3), 727–747.
[15] Avramov, D., & Wermers, R. (2006). Investing in mutual funds when returns are predictable. *Journal of Financial Economics, 81*(2), 339–377.