ADVANCE SOCIAL SCIENCE ARCHIVE JOURNAL
Available Online: https://assajournal.com
Vol. 05 No. 01. Jan-March 2026.Page#5012-5032
Print ISSN: 3006-2497 Online ISSN: 3006-2500
Platform & Workflow by: Open Journal Systems
https://doi.org/10.5281/zenodo.21837945

Multidimensional Timing Ability in Pension Fund Management: Evidence from Conventional and Islamic Funds

Ghulam Nabi (Corresponding Author)
P. hD, College of Statistical Sciences, University of the Punjab, Quaid-i-Azam Campus, Lahore 54590, Pakistan.
gnabi.stat@pu.edu.pk

ABSTRACT
This study examines whether pension fund managers have multidimensional timing ability or merely raise the position in the equity portfolio prior to positive market returns. There are 316,448 daily observations for 69 Pakistani pension funds that span from June 2007 to July 2026. For this analysis, benchmark-aligned, the fund-day observation covers 61,520 observations (from 10 January 2014 to 10 July 2026) of 24 equity sub-funds, comprising 11 conventional funds and 13 Islamic funds, matched to the KSE-100 Total Return Index. The Treynor–Mazuy and Henriksson–Merton models are linked with an integrated specification where beta is dependent on market volatility, the NAV-implied liquidity and valuation-friction state, skewness and kurtosis. The fund-level estimates are robustified by equal-weight portfolios, regime analysis, alternative rolling windows, and a machine-learning holdout for the period July 2023 to July 2026. The findings point to a lack of the classical market time-taking ability but suggest that the defensive behaviour is heterogeneous, specifically volatility timing, liquidity timing and co-skewness timing in particular funds. Islamic funds are more volatile, but there is no disadvantage in terms of persistent return, return to volatility or drawdown. Adaptability in forecasting models is not always profitable with respect to the historical-mean benchmark, showing that state-dependent behaviour is not even exploitable.

Keywords*: Pension Funds; Market Timing; Volatility Timing; Liquidity Timing; Higher-Moment Timing; Islamic Finance; Machine Learning; Pakistan*
JEL Classification: G11; G12; G23; C52; C53

1. Introduction

Often, the only question asked about investment-management skill is whether the manager beat the market. The economic question is more wide-ranging for pension investors. A manager can help by investing in securities with a higher level of security selection, avoiding illiquid states, avoiding extreme things that can happen in the left tail, buying into favourable skewness, or defensive beta reductions when volatility increases. An apparently favourable timing coefficient may also be due to cash flows, non-linear relations between benchmarks, out-of-date prices, and option-like positions, mismatch of benchmarks or sampling error. The difference is important, because retirement plans are long horizon contracts where downside risks, sequence-of-returns risk, liquidity and constant fees should be more important than the amount of a small difference in average return.

Pakistan is a good resource for information. The Voluntary Pension System (VPS) is an individually funded, portable and defined contribution pension plan (MUFAP, 2026). The Securities and Exchange Commission of Pakistan (SECP) expanded the framework in 2024 to allow fully funded employer- and government-sponsored schemes through licensed institutions

(SECP, 2024a, 2024b). KPK had already shifted the procedure of civil employees appointed from 7 June 2022 towards the contributory system (SECP, 2023). These reforms make more important the role of credible measures of performance; poor measures might lead to incentives for statistical noise, while good measures could help determine whether managers adjust exposure and do so in a manner that benefits retirement outcomes.

The empirical literature is not unanimous in regard to the timing skill. Treynor and Mazuy (1966) investigated the convexity of fund returns in market returns, and Henriksson and Merton (1981) explored the question of whether managers switch between up-market and down-market betas. Generally, daily data are more powerful than monthly data (Bollen & Busse, 2001), and the nonlinear holdings, dynamic portfolios, choice of benchmarks, and trading mechanics can produce misleading curvature (Jagannathan & Korajczyk, 1986; Busse et al., 2023). Research has thus shifted to multi-dimensional and conditional timing. Volatility timing (Busse, 1999), liquidity timing (Cao et al., 2013; Foran & O'Sullivan, 2017), selectivity, market timing, and volatility timing (Ferson & Mo, 2016), dynamic relationships (Wattanatorn & Padungsaksawasdi, 2021), and higher order co-moments (Busse et al., 2024). The common perception would be that managers could be wrong about the direction of the market, but perhaps they can control the risk in a condition.

This study takes this view to conventional and Islamic pension equity funds. Screening for Shariah is carried out to exclude businesses that are prohibited, limit leverage, interest-based instruments and some hedging. There are potential constraints which may limit diversification and tactical flexibility, but may also minimize exposure to highly leveraged or financially fragile firms. In theory, then, the answer to that question is unclear: Islamic funds could have fewer instruments available to them with which to change beta, or they may have a structurally more defensive screened opportunity set. The evidence from around the world is inconclusive, especially in markets outside the major financial hubs (Mohammad & Ashraf, 2015; Delle Foglie & Panetta, 2020; Alqahtani et al., 2025). A direct comparison becomes even more relevant with recent evidence from Pakistan, which indicates that asset allocation is influenced by the governance and Shariah structures (Khaleel et al., 2026).

The following questions are examined in this study (1) What is the evidence of positive classical market timing for conventional and Islamic pension equity managers? (2) Is beta a constant for a company? (3) Do the pension fund managers expand exposures in more liquid market and valuation states? (4) Do pension fund managers increase exposure when the market is skewed and lagged, and do they decrease exposure when the excess kurtosis is high? (5) Can machine-learning models predict the out-of-sample performance of the next month better if it is adjusted for the benchmark?

Signs that are collected at the fund level are differentiated from signs that are collected in the industry. Newey–West standard errors correct for serial correlation and heteroskedasticity, and Benjamini-Hochberg false-discovery-rate adjustment prevents the misinterpretation of chance related findings as skill. Conventional and Islamic portfolios using equal weights are used to determine if a pattern persists at the system level; median-split regimes are used to test performance in high-volatility, low-liquidity, and joint-stress regimes. The forecasting design is also very chronological: models are trained up till 2022, and then test their performance from 2023 to July 2026, for which there will be no flexibility, as the reward will only come when the model's flexibility is useful in truly unseen forecasting (Campbell & Thompson, 2008; Gu et al., 2020).

The descriptive evidence does not provide evidence of performance dominance. The mean annual returns of conventional and Islamic funds are 19.94% and 22.39% respectively with statistically insignificant difference. Islamic funds suffer from considerably higher volatility (20.00% versus 18.60%) but the systems are not significantly different with respect to return-to-volatility ratios, skewness, excess kurtosis, or maximum draw-down. The timing coefficients for classical funds are mostly negative and the integrated market timing coefficient is even more negative for Islamic funds. This means that the performance of the Islamic funds after taking conditional states into account is less convex compared to the common KSE-100 index.

Conditional states give rise to a more complex pattern. Conventional funds have negative mean timing of volatility, positive mean timing of liquidity and positive mean timing of co-skewness. There are no significant estimates for the 11 conventional funds, and co-skewness is positive for all of them. Islamic funds also exhibit positive median characteristics in a number of areas, though the impact is not as consistent. The portfolio coefficients of the different managers are not statistically different, so are not reflections of industry-wide skill. This interpretation is confirmed by regime results: both systems are found to be good in liquid state and bad in the case of high volatility and low liquidity.

Results from machine learning augment an extra discipline. There is no more flexible model than the historical-mean model. The model's out-of-sample R^2 values are all negative, and the benchmark of historical means is set to be zero; directional accuracy is below 46%. The active-return positive observations go down from 52.9% to 41.0% of the total in test. The results suggest regime instability and add to the separation of historical conditional behaviour and tradable predictability.

The institutional and conceptual context is in Section 2. The literature is reviewed and hypotheses are developed in Section 3. In Section 4, the data, variables, models, inference, robustness tests and forecasting design are described. The empirical results are presented in Section 5, followed by a discussion of the contributions and implications in Section 6, and a conclusion in Section 7.

2. Institutional Setting and Conceptual Framework

2.1 Pakistan's defined-contribution pension architecture

The VPS was created in Pakistan as an alternative to the traditional "pay-as-you-go" retirement promises that are not funded. A participant spreads out investments in sub-funds, usually equity, debt, money market and commodity/gold, and the value of retirement wealth will be dependent on market performance, fees, the manner of the allocation, and the quality of professional management. The portability and individual ownership of the framework mean that the performance of the NAV is directly relevant for the beneficiaries (MUFAP, 2026). The 2024 amendments to the regulations broadened the definition of employer-sponsored and government-sponsored defined-contribution plans and focused on the professional management of plans by licensed entities (SECP, 2024a, 2024b). This means an unbalanced panel and creates differences in fund age due to the voluntary schemes which are older and those which are employer and provincial and are newer.

There are three distinguishing characteristics of pension sub-funds as compared to conventional retail-type mutual funds. One is that they have a long horizon, so drawdowns and sequence-of-returns risk are significant for retirement consumption. Second, policy and lifecycle allocation can impact contributions and withdrawals, not performance chasing. Third, the strategic roles of sub funds vary: Equity funds offer systematic growth exposure; debt funds offer a stable wealth, whereas commodity/money market funds offer diversification of risk. The timing analysis is limited to equity sub-funds due to the different benchmarks as well as state variables used in these sub-funds.

2.2 Why conventional and Islamic timing may differ

A traditional equity manager can choose from the entire investable universe allowed in the mandate. An Islamic manager works in an Islamic universe and its universe is screened through Shariah, which means that there are a number of constraints applied to the universe that the Islamic manager works in. A more restricted opportunity set can help to limit tactical flexibility, especially in down markets, and can help limit the number of instruments available for quick beta change.

There is another prediction that is equally possible. Keeping conventional banks and highly leveraged companies out of the group could help lower the risk of financial fragility and left-tail risk. When an active forecast is not performed, an Islamic fund can thus seem more resilient. This is a very important distinction because it is not managerial timing. State-dependent changes in exposure are required to achieve timing; stable exposures in the portfolio are required to achieve structural resilience.

2.3 A multidimensional view of timing

Let a fund's conditional market beta be $\beta_{i,t}$. A multidimensional manager may choose beta using forecasts of and observable risk states$r_{m,t+1}$:

$$\beta_{i,t} = g\left(E_t r_{m,t+1}, E_t \sigma_{m,t+1}, E_t L_{t+1}, E_t S_{m,t+1}, E_t K_{m,t+1}, I_t\right)$$

Here, σ represents market volatility, L is liquidity, S is skewness, K is kurtosis and I_t is the set of information the manager has. Return timing suggests that good market returns are preceded by a high beta. Volatility timing suggests that one should have a lower beta when they have higher expected risk and liquidity timing suggests that they should be more exposed when trading and valuation conditions are more favourable. Co-skewness timing is favourable when the state has positive asymmetry, while co-kurtosis timing is the opposite effect, i.e., the thicker the tails the lower the exposure is. The resulting responses may cancel each other out, and a one-dimensional return-timing test may be inadequate to uncover economically rational risk management.

The difference also aids in performance measurement. The alpha is average return that cannot be attributed to included risk exposures and the timing coefficients are the changes in exposure based on market states. A positive alpha is not a sign of successful timing, and success in timing doesn't necessarily result in a positive net performance after fees. Governance of pension needs, therefore, must be demonstrated both in an unconditional and conditional way.

3. Literature Review and Hypothesis Development

3.1 From unconditional performance to market timing

Modern performance evaluation separates security-selection skill from systematic risk exposure. Jensen (1968) defines alpha as return beyond that predicted by a market model, while Carhart (1997) shows that common style factors and momentum explain part of apparent persistence. Fama and French (2010) further demonstrate that the mutual-fund industry, after costs, rarely earns persistent benchmark-adjusted returns. The remaining question is whether managers can create value by changing risk exposure at appropriate times even when unconditional alpha is small.

Treynor and Mazuy (1966) suggest that convexity exists between fund and market returns when they are successful timers. A quadratic market term should be positive, if market gains when manager raises beta and market loses when the manager lowers beta. Henriksson and Merton (1981) use the concept of timing as a switch between market regimes in the form of an option. Both models are still very common due to being parsimonious and interpretable. However, the track record is not entirely impressive. There is little systematic timing found by Chang and Lewellen (1984) and many other studies, and Bollen and Busse (2001) demonstrate that more evidence of timing is seen with daily observations as compared to monthly observations, due to aggregation reducing statistical power. Jiang (2003) formulates a nonparametric test, and is also aware that timing conclusions may be model dependent.

A negative timing coefficient does not automatically establish managerial incompetence. Option-like securities, nonlinear payoffs, cash balances, investor flows, stale prices, and benchmark dynamics can create artificial timing patterns (Jagannathan & Korajczyk, 1986; Busse et al., 2023). In a large cross-section, some significant estimates will also arise by chance (Kosowski et al., 2006; Barras et al., 2010). Prior to interpreting fund-level coefficients as skill, it is therefore important to have robust covariance estimates and false-discovery control.

Hence, the first is a conservative one:

H1: Pension equity funds on average do not show up robust positive timing after multiple-testing adjustment, according to a Treynor–Mazuy or Henriksson–Merton test.

This null type hypothesis supports the evidence found internationally, and has set the expectation high for skills to be claimed. Rejection needs more than just good sample means; it needs to have positive coefficients that are robustly estimated.

3.2 Volatility timing as defensive risk management

Volatility timing is predicated on the idea that, when market returns are difficult to predict, a manager can be useful by adjusting to the conditional variance of the market. Volatility in mutual funds is reported by Busse (1999) to be related to market volatility and to the behavior of mutual-fund managers during turbulent times, which is captured by daily-return tests. The economic rationale is simple. With positive risk aversion, a lower beta in high-volatility states may enhance expected utility despite a decrease in one-for-one expected return. With a positive risk aversion, the lower beta in high volatility states could lead to higher expected utility and lower drawdowns. At the strategy level, Moreira and Muir (2017) demonstrate that volatility-managed portfolios can enhance risk-adjusted performance by adjusting exposure inversely to the volatility of the past, but that it depends on the strategy's implementation choices, turnover and estimation choices.

In a stochastic-discount-factor framework, Ferson and Mo (2016) incorporate volatility timing to simultaneously select and time. Foran and O'Sullivan (2017) conclude that it is important to consider return, volatility and liquidity timing simultaneously due to the fact that a coefficient can accommodate another omitted timing channel. The insight of the joint estimation is particularly important to pension funds. A manager might seem to have a concave return relationship – he/she may short out of the market during periods of heightened volatility – yet protect long horizon beneficiaries from adverse variance.

H2: Pension managers exhibit defensive volatility timing, reflected in a negative beta response to lagged market volatility, although the effect need not be uniform across funds.

The hypothesis does not make any a priori distinction between Islamic and conventional funds. While Shariah restrictions can reduce tactical flexibility, structural screening can help decrease the reliance on active beta cuts.

3.3 Liquidity timing and valuation friction

Liquidity is a state variable since the price effect and execution cost of portfolio changes can be different at different times. Aggregate liquidity risk is priced (Pastor and Stambaugh 2003) and a popular return-to-volume illiquidity measure is developed (Amihud 2002). Before market liquid states, do the mutual-fund managers time market liquidity? Cao et al. (2013) question this and determine that some do. Volatility and liquidity timing are not the same thing, according to Foran and O'Sullivan (2017). In less mature markets, being able to adjust through anticipation may be especially critical due to the high costs of adjustment in the form of trading costs, price limits, concentrated ownership, and occasional order flow.

A favourable liquidity timer will have a higher beta where the lagged state corresponds to a lower number of zero returns and lower dispersion. The proxy may also be reflective of operational NAV practices and is therefore interpreted as market liquidity and efficiency of valuation. More recently, there has been some evidence that mutual-fund timing can be sensitive to the liquidity measure employed (Yin et al., 2024), which further underscores this caution.

H3: Pension managers increase market exposure in more favourable NAV-implied liquidity and valuation states, producing positive liquidity-timing coefficients and stronger active returns in high-liquidity regimes.

3.4 Co-skewness, co-kurtosis, and tail-state timing

The mean-variance analysis is not sufficient when returns are asymmetric and fat tailed. Conditional skewness is extended to asset pricing by Kraus and Litzenberger (1976) and conditional skewness helps to explain expected returns by Harvey and Siddique (2000). Positive skewness tends to favour investors, as it increases the likelihood of big rewards, while systematic negative skewness is a dislike. Excessive co-skewness and co-kurtosis can be risk premia as shown by Fang and Lai (1997). The fourth moment is especially important for a pension portfolio because of the interaction with withdrawal and contribution horizons and because extreme losses can affect the pension portfolio.

Higher order timing is not just about keeping unconditional securities with great moments. A co-skewness timer switches the market exposure when the systematic distribution becomes more or less skewed. A co-kurtosis or tail risk timer is designed to hedge against the risk of extreme outcomes. Wattanatorn and Padungsaksawasdi (2021) suggest a higher order co-moment timing model and identify that superior performing mutual funds can benefit from co-kurtosis states. This is a small and limited body of literature, particularly within the fields of pensions, Islamic funds and emerging markets. This paucity provides a strong possibility of "contribution" but also a warning against mechanical interpretation of polynomial coefficients. Cubic and quartic market terms are very collinear and outlier sensitive. It is therefore appropriate to look at the excess kurtosis and lagged skewness of the rolling market and engage them linearly in the current market return, as is done in the present article. This leaves the interpretation as a state dependent beta response.

Positive sign values are based on preference for skewness timing and negative sign values are based on preference for kurtosis timing. However, kurtosis is unsigned; a high kurtosis value means that tails are thick but not whether the extreme next move will be a positive one or a negative one. Thus, co-kurtosis timing is not about direction, but about avoiding tails.

H4a: The co-skewness-timing coefficient is positive because managers increase beta when lagged market skewness is more favourable.

H4b: The co-kurtosis-timing coefficient is negative because managers reduce beta when lagged excess kurtosis, and therefore tail risk, is high.

3.5 Islamic versus conventional fund performance

The comparative Islamic-fund literature does not support a universal ranking. Shariah screens may increase concentration and tracking error, but they may also reduce leverage and exposure to conventional financial firms. Performance therefore depends on market conditions and benchmark choice (Mohammad & Ashraf, 2015). There is no conclusive evidence from systematic reviews of greater system superiority (Delle Foglie & Panetta, 2020), and more recent evidence indicates that relative performance may vary across extreme events (Alqahtani et al., 2025).

Pakistan is a useful comparison since the conventional and Islamic pension schemes are set in a similar overarching regulatory framework. Shariah governance has an impact on diversification and asset allocation (Khaleel et al., 2026), so there is a common benchmark, which is the broad-market opportunity cost, but may penalise the sector exclusions in the structure. The use of official KSE-100 Total Return Index is done to make the study comparable; a parallel study on the KMI-30 is an important robustness extension.

H5: The overall difference in the net risk-adjusted returns between conventional and Islamic pension funds holds no persistent statistical evidence of difference whatsoever, but there could be differences in the timing and volatility channels.

3.6 Machine learning, prediction, and the danger of flexible overfitting

There are three types of leakage in fund forecasting: train-test splits may leak future regimes into model tuning; contemporaneous variables may be misinterpreted as predictors; and the target may be mechanically contained in an input. A valid design should be: chronological, based only on information known at the forecast time, and compare each model against a simple model. Campbell and Thompson (2008) examine if forecasts beat the historical mean in terms of squared-error loss; Diebold and Mariano (1995) offer formal logic for comparing forecasts.

The econometric timing model is not replaced by machine learning, but rather it is used as a validation exercise. If the flexible algorithms are not able to predict what the benchmark adjusted performance will be in an unseen timeframe, the state coeffs are intended to be descriptive of the historical conditional behaviour and not to be trading rules. If it is a consistent out-of-sample gain, it would give more credibility to there being information to be exploited.

4. Data and Methodology

4.1 Data sources and sample construction

The fund panel consists of two workbooks for the study: conventional-bank pension funds and Islamic-bank pension funds. The information on the daily records is the sector, asset management company, fund name, sub-fund type, inception date, NAV, total expense ratio (monthly and yearly), management fee, sales and marketing charge and validity date. The workbook with traditional format includes 147,749 observations over a total of 29 named funds and 55 fund-category series from July 2007 to July 2026. There are 168,699 observations for 40 named funds and 72 series over the period June 2007 to July 2026 in the Islamic workbook. There are 316,448 observations in the files and 69 named funds and 127 series.

The benchmark used is the official Pakistan Stock Exchange (PSX) KSE-30 Total Return Index which has been prepared from PSX historical data from 2014-2026. The index is free-float weighted and includes distributions so conventional timing coefficients can be compared to a broad-market index based on Islamic timing. However, there is the possibility of mismatch with the Shariah opportunity set, since it is not replicated. Thus, the dual-benchmark extension with the KMI-30 Total Return Index is identified for future research.

The analysis applies only to equity sub-funds as debt, money-market and commodity funds have different benchmarks and state variables. A fund is retained when at least 500 valid daily observations can be aligned with PSX trading dates. The final benchmark-matched sample contains 24 funds—11 conventional and 13 Islamic—with 61,520 fund-day observations from 10 January 2014 to 10 July 2026. Table 1 distinguishes this analytical sample from the broader source panel covering June 2007 to July 2026.

Table 1. Sample construction and coverage

Dataset	Daily observations	Number of funds	Fund-category series	Coverage
Conventional Fund	147,749	29	55	Jul 2007-Jul 2026
Islamic Funds	168,699	40	72	Jun 2007-Jul 2026
Combined Funds	316,448	69	127	Jun 2007-Jul 2026
Equity timing sample	61,520	24	24	Jan 2014-Jul 2026

Source: Authors' calculations from the supplied pension-fund NAV panel and PSX data.

4.2 Data cleaning and return construction

The raw panel has 5,747 non-positive NAV observations, around 1.82% of all the observations, 73 of which are negative observations in the conventional file. Logarithmic returns are not valid if the NAV is non-positive, nor do such returns make economic sense in the context of typical open-end fund accounting. They do not get replaced but are set to missing. The daily return on the funds is:

$$r_{i,t} = \ln\left(\frac{NAV_{i,t}}{NAV_{i,t-1}}\right) \quad (1)$$

and benchmark return is

$$r_{m,t} = \ln\left(\frac{P_{m,t}}{P_{m,t-1}}\right) \quad (2)$$

Here $P_{m,t}$ is the KSE-100 Total Return Index. Observations with $|r_{m,t}| > 0.25$ are treated as likely data errors or unadjusted unit changes/distributions. Winsorization is performed at the 0.25th and 99.75th percentile of the analytic sample. The narrow tail treatment allows keeping true crisis variation whilst reducing the impact of extreme NAV anomalies on squared and interaction terms.

The timing models do not employ excess returns, because a comprehensive series of consistently dated daily risk-free returns was not combined into the current analysis. This decision makes a big difference to the interpretation of alpha and a little difference to beta. The quadratic coefficient and state-interaction coefficient are not mathematically invariant, although they are less sensitive since the risk-free rate is smooth with daily frequency relative to equity returns. The article does not, therefore, classify the intercept as a definite Jensen alpha. A future version should incorporate State Bank of Pakistan Treasury-bill data/repo data and should repeat all of the models in excess-return terms.

Expense fields are not considered as primary regressors as their raw distributions contain implausible positive and negative values, and due to different conventions of reporting. The returns are based on NAV, so include the expenses the fund incurs if they are "embedded" in NAV. The separate paper which would cover fee determinants would need to be validated against fund reports and a consistent annual expense measure.

4.3 Descriptive performance measures

For each fund, annualized arithmetic mean log return and volatility are calculated as

$$r_i^{ann} = 252\bar{r}_i, \quad \sigma_i^{ann} = \sqrt{252}\, s(r_i) \quad (3)$$

The return-to-volatility ratio is r_i^{ann}/σ_i^{ann}. It is reported rather than called a Sharpe ratio because the numerator is not an excess return. Sample skewness and excess kurtosis describe non-normality. Maximum drawdown is calculated from the cumulative wealth index $W_{i,t} = \exp\left(\sum_{\tau=1}^{t} r_{i,\tau}\right)$ as

$$MDD_i = \min_t\left(\frac{W_{i,t}}{\max_{\tau\le t} W_{i,\tau}} - 1\right) \quad (4)$$

System-level descriptive statistics are equal-weight averages across funds so that a large or older scheme does not dominate. Cross-system differences are evaluated with the Mann-Whitney test because the number of funds is small and coefficient distributions are non-normal.

4.4 Classical market-timing models

The Treynor–Mazuy model is

$$r_{i,t} = \alpha_i + \beta_i r_{m,t} + \gamma_{TM,i} r_{m,t}^2 + \varepsilon_{i,t} \quad (5)$$

A positive $\gamma_{TM,i}$ indicates convexity consistent with successful return timing. The Henriksson–Merton model is

$$r_{i,t} = \alpha_i + \beta_i r_{m,t} + \gamma_{HM,i} \max\left(0, r_{m,t}\right) + \varepsilon_{i,t} \, (6)$$

A positive $\gamma_{HM,i}$ indicates that beta is higher in up markets. Equation (6) is a return-based implementation of the directional timing idea; the interaction can also be written with a binary up-market indicator. Both models are estimated separately for each fund with Newey–West HAC standard errors using five lags.

4.5 Conditional volatility, skewness, and kurtosis states

Market states are estimated from information available before the current return. Using a 63-trading-day rolling window, approximately one quarter, the annualized volatility state is

$$\sigma_{m,t-1} = \left[\frac{252}{62} \sum_{j=1}^{63} \left(r_{m,t-j} - \bar{r}_m\right)^2\right]^{1/2} \qquad (7)$$

Rolling skewness and excess kurtosis are calculated from central moments, where $\mu_{q,t-1}$ denotes the qth central moment of the previous 63 market returns.

$$S_{m,t-1} = \frac{\mu_{3,t-1}}{\mu_{2,t-1}^{3/2}} (8)$$

$$K_{m,t-1} = \frac{\mu_{4,t-1}}{\mu_{2,t-1}^{2}} - 3 \quad (9)$$

Each state is standardized over the available benchmark history:

$$z(X_{t-1}) = \frac{X_{t-1} - \bar{X}}{s_X} \qquad (10)$$

The one-day lag prevents contemporaneous market outcomes from entering the manager's information set. Robustness tests replace 63 days with 42 and 126 days.

4.6 NAV-implied liquidity and valuation-friction state

Suppose that for each trading day t, the number of fund returns being observed is N_t, the proportion of fund returns with absolute value less than or equal to a small number is Z_t, and the cross-sectional median absolute deviation (MAD) of fund returns is D_t:

$$Z_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \mathbf{1}\left(\left|r_{i,t}\right| < 10^{-8}\right) \qquad (11)$$

$$D_t = \text{median}_i \left|r_{i,t} - \text{median}_j\left(r_{j,t}\right)\right| \quad (12)$$

The liquidity state is

$$L_t = -\frac{1}{2}\left[z(Z_t) + z(D_t)\right] \qquad (13)$$

where z(•) means standardization. The L_t is also high, which translates to less dispersion in the NAV values across the cross-section (fewer zero returns) and better liquidity. L_{t-1} is the variable used in regressions. While equation (13) represents a methodological contribution for cases where holdings data is not available, it is not an assertion that fund NAVs reveal exchange bid-ask spreads. The lack of returns can be due to out-of-date prices, synchronised holidays, low volatility or revaluation. That is why the article triangulates the coefficient with the outcome of the regimes and labels the construct consistently.

4.7 Unified multidimensional timing model

The central empirical specification is

$$r_{i,t} = \alpha_i + \beta_i r_{m,t} + \gamma_{M,i} r_{m,t}^2 + \gamma_{V,i} r_{m,t} \sigma_{m,t-1} + \gamma_{L,i} r_{m,t} L_{t-1} + \gamma_{S,i} r_{m,t} S_{m,t-1} + \gamma_{K,i} r_{m,t} K_{m,t-1} + \varepsilon_{i,t} \quad (14)$$

The model interprets timing as conditional beta adjustment. The instantaneous beta implied by Equation (14) is

$$\beta_{i,t}^* = \beta_i + 2\gamma_{M,i} r_{m,t} + \gamma_{V,i} \sigma_{m,t-1} + \gamma_{L,i} L_{t-1} + \gamma_{S,i} S_{m,t-1} + \gamma_{K,i} K_{m,t-1} \quad (15)$$

Favorable signs are $\gamma_{M,i} > 0$, $\gamma_{V,i} < 0$, $\gamma_{L,i} > 0$, $\gamma_{S,i} > 0$ and $\gamma_{K,i} < 0$. Fund by fund estimation of equation (14) uses Newey-West standard errors and ten lags, due to the persistence of state interactions. The joint model avoids omitted-channel bias: volatility and liquidity (and higher moments) are included jointly.

4.8 Multiple testing and aggregate portfolios

The standard 5% tests would lead to false discoveries with 24 funds and a number of coefficients. For each of the timing dimensions, the fund-level p-values are adjusted with Benjamini-Hochberg (1995) false-discovery-rate procedure. If $p_{(1)} \leq \ldots \leq p_{(k)}$ are ordered p-values, the largest k satisfying

$$p_{(k)} \leq \frac{k}{m} q \quad (16)$$

The target false-discovery rate is q = 0.05 and up to k hypotheses are rejected. Only a fund with a coefficient that has the economically favorable sign and an adjusted p-value below 0.05 is considered significantly favorable.

To check if the timing is an industry-wide property equal-weights conventional and Islamic daily portfolios are created with the help of at least three funds. Equation (14) is then estimated for each system portfolio. This aggregation reduces idiosyncratic NAV noise but can hide heterogeneous manager skill. Fund-level and aggregate results are therefore interpreted jointly.

4.9 Volatility-liquidity regimes

Days are classified using medians of lagged volatility and lagged liquidity. The main regimes are low versus high volatility, low versus high liquidity, low-volatility/high-liquidity and high-volatility/low-liquidity. For system g, annualized active return in regime R is

$$AR_{g,R} = 252\,\mathrm{mean}\left(r_{g,t} - r_{m,t} \mid t \in R\right) \quad (17)$$

A market beta is estimated within each regime.

4.10 Machine-learning validation

Daily observations are converted to a monthly fund panel. The prediction target is next-month benchmark-adjusted return:

$$y_{i,t+1} = r_{i,t+1}^{month} - r_{m,t+1}^{month} \quad (18)$$

All predictors are known at month t: one- and three-month fund and benchmark returns, rolling fund volatility, downside volatility, skewness, kurtosis, zero-return share, market volatility, market skewness, market kurtosis, system indicator and fund identity. Continuous predictors are standardized within the training data for penalized linear models. Categorical identifiers are one-hot encoded. The models are historical mean, Ridge, Elastic Net, Random Forest, Extra Trees and Gradient Boosting. Hyperparameters are intentionally moderate to reduce the risk of fitting a small monthly panel.

out-of-sample R_{OS}^2:

$$R_{OS}^2 = 1 - \frac{\sum_{test}\left(y_{i,t} - \hat{y}_{i,t}\right)^2}{\sum_{test}\left(y_{i,t} - \bar{y}_{train}\right)^2} \quad (19)$$

A positive R_{OS}^2 indicates improvement over the training-sample historical mean, a negative value indicates deterioration, and the historical-mean benchmark itself equals zero by construction. The exercise predicts relative performance rather than the market index and does not assume that pension funds are freely tradable portfolios.

4.11 Robustness tests

Four robustness strategies are employed. First, market states are computed over 42, 63, and 126 trading days. Secondly, the acute COVID-19 interval is omitted to test if one crisis is sufficient to push the coefficients. Thirdly, the classical TM and HM estimates are compared to the integrated model. Fourth, the fund-level distributions are compared to an equal-weight aggregate portfolio. Robustness is not an attempt to find significance but rather a tougher sign of greater uncertainty or a single low p-value is more convincing.

5. Empirical Results

5.1 Descriptive performance and risk

Table 2 reports equal-weight averages of fund-level performance measures. Conventional and Islamic equity funds earn annualized mean returns of 19.94% and 22.39%, respectively; the 2.45 percentage-point difference is not statistically significant. Islamic funds have significantly higher annualized volatility (20.00% versus 18.60%; Mann-Whitney p = 0.0065), but the return-to-volatility ratio, skewness, excess kurtosis, and maximum drawdown do not differ significantly. Both systems have negative skewness and high excess kurtosis demonstrating the inadequacy of using mean-variance comparisons. H5 therefore receives mixed support: broad performance measures show no persistent system gap, but volatility differs significantly.

Table 2. Descriptive performance of the equity timing sample

System	Funds	Fund-days	Annual return (%)	Annual volatility (%)	Return/volatility	Skewness	Excess kurtosis	Max drawdown (%)
Conventional	11	28,749	19.94	18.60	1.06	-0.22	3.54	-38.50
Islamic	13	32,771	22.39	20.00	1.11	-0.18	4.11	-35.71

Source: Authors' calculations. Returns are annualized daily log returns; the ratio is not a risk-free-rate-adjusted Sharpe ratio.

As shown, both systems have negative skewness and a large excess kurtosis, which is why a mean-variance comparison is not sufficient. The downsides are asymmetric, and large observations happen more than they do in a Gaussian distribution. The negative average skewness and mean maximum drawdown of Islamic funds are a bit lower than those of conventional funds, however both differences are not statistically significant. H5 that Sharia classification is associated with higher volatility in this sample is supported.

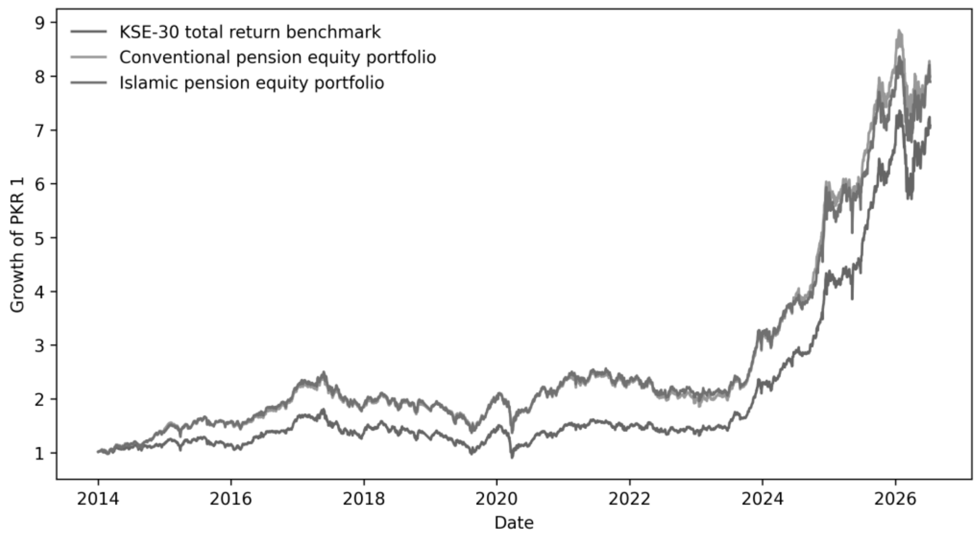

Figure 1. Cumulative wealth of KSE-100 and equal-weight pension equity systems.

Table 3. Cross-system tests of performance and timing coefficients

Measure	Conventional mean	Islamic mean	Islamic minus conventional	Mann-Whitney p-value
Annual return	0.1994	0.2239	0.0245	1.0000
Annual volatility	0.1860	0.2000	0.0140	0.0065
Return/volatility	1.0552	1.1111	0.0559	0.9538
Maximum drawdown	-0.3850	-0.3571	0.0279	0.8167
Treynor–Mazuy timing	-0.3947	-0.5755	-0.1808	0.0725
Henriksson–Merton timing	-0.0747	-0.0966	-0.0219	0.0426
Integrated market timing	-0.3088	-0.6339	-0.3251	0.0021
Volatility timing	-0.0203	0.0016	0.0219	0.3539
NAV-implied liquidity timing	0.0239	0.0060	-0.0179	0.5239
Co-skewness timing	0.0271	0.0087	-0.0185	0.0010
Co-kurtosis timing	0.0055	0.0043	-0.0013	0.2970

Source: Authors' calculations. P-values compare fund-level distributions and do not adjust the displayed row set for multiple comparisons.

5.2 Classical and integrated timing ability

Positive market timing is not supported by the classical models. The traditional mean Treynor–Mazuy coefficient has a negative value of -0.395, and one out of 11 funds has a positive value. The Islamic mean is -0.575 and none of 13 funds has a favorable sign. Negative Henriksson-Merton coefficients for all the conventional funds and for all the Islamic funds. There is no positive classical timing coefficient after false-discovery adjustment. This means that H1 is supported.

The integrated market-timing coefficient remains negative: -0.309 for conventional funds and -0.634 for Islamic funds. The Islamic distribution is significantly more negative (Mann-Whitney p = 0.0021). This disparity might be related to reduced tactical flexibility, sector exposures, benchmark mismatch or increased participation in negative curvature. It should not be taken as an indication of poorer average performance as there is no significant return or return-to-volatility disadvantage as indicated in Table 2. What is realized, however, is that the returns of Islamic funds are less convex relative to the common KSE-30 benchmark even after controlling the conditional states.

Table 4 reveals the richer timing pattern. Favorable volatility timing is present in 63.6% of conventional and 61.5% of Islamic funds, based on the expected negative sign, although only one conventional fund survives FDR adjustment. Liquidity-timing signs are favorable for 90.9% of conventional funds and 76.9% of Islamic funds. Co-skewness timing is positive for all conventional funds and 84.6% of Islamic funds; one conventional estimate is FDR significant. Co-kurtosis evidence is weaker and unstable: only 36.4% of conventional funds and 61.5% of Islamic funds have the favorable negative sign. Thus H2, H3, and H4a receive sign-based but not industry-wide statistical support; H4b receives little support.

Table 4. Fund-level multidimensional timing summary

System	Timing dimension	Favorable sign	Mean coefficient	Median coefficient	Favorable funds (%)	FDR-significant favorable funds
Conventional	Treynor–Mazuy market	+	-0.3947	-0.3866	9.1	0
Conventional	Henriksson–Merton market	+	-0.0747	-0.0668	0.0	0
Conventional	Integrated market	+	-0.3088	-0.3469	9.1	0
Conventional	Volatility	-	-0.0203	-0.0092	63.6	1
Conventional	NAV-implied liquidity	+	0.0239	0.0242	90.9	0
Conventional	Co-skewness	+	0.0271	0.0232	100.0	1
Conventional	Co-kurtosis/tail risk	-	0.0055	0.0020	36.4	0
Islamic	Treynor–Mazuy market	+	-0.5755	-0.6116	0.0	0
Islamic	Henriksson–Merton market	+	-0.0966	-0.1004	0.0	0
Islamic	Integrated market	+	-0.6339	-0.6660	0.0	0
Islamic	Volatility	-	0.0016	-0.0038	61.5	0
Islamic	NAV-implied liquidity	+	0.0060	0.0169	76.9	0
Islamic	Co-skewness	+	0.0087	0.0102	84.6	0
Islamic	Co-kurtosis/tail risk	-	0.0043	-0.0027	61.5	0

Source: Authors' estimates from Equation (14). FDR significance uses Benjamini-Hochberg adjustment within each timing dimension.

The mean state-timing coefficients with cross-sectional 95% confidence intervals are shown in figure 2. The conventional volatility coefficient is on the good side (in the negative sense), as are liquidity and co-skewness. Islamic signs are directionally positive at the median but close to zero on average. The intervals highlight differences and avoid overemphasizing accuracy.

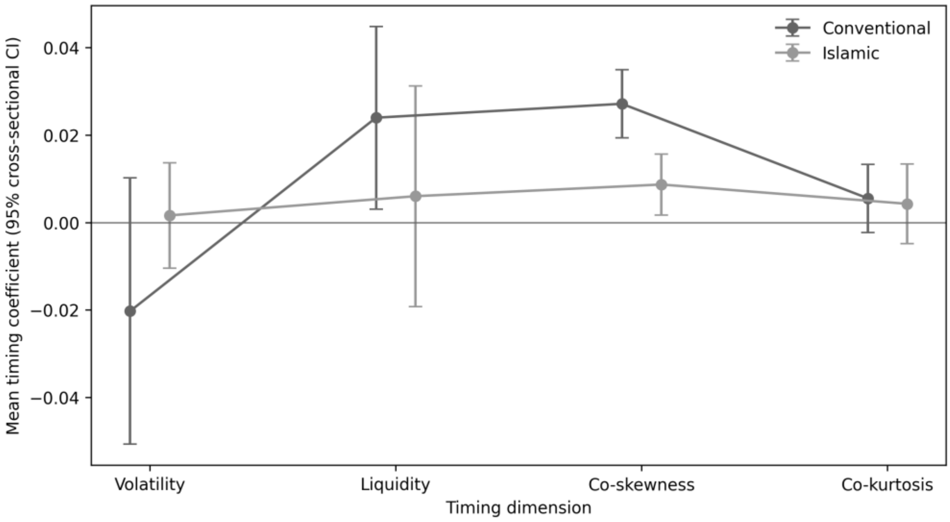

Figure 2. Mean multidimensional timing coefficients and cross-sectional 95% intervals.

5.3 Aggregate system portfolios

The value of Equation (14) with equal-weight system returns is estimated using Table 5. Conventional and Islamic funds have Market betas of 0.873 and 0.902, respectively, with both well estimated. The slightly higher Islamic beta is in line with the higher annual volatility. Intercept is positive, but small. The market-timing coefficients of the aggregate are still negative and small. Conventional volatility timing is negative, but liquidity and co-skewness timing is positive and the p-values are 0.277, 0.163 and 0.105 respectively. Liquidity timing based on Islamic principles is also positive but negligible. These results will be useful for calibration; good cross-sectional results are not evidence for a systematic capability.

Table 5. Aggregate equal-weight integrated timing models

System	Term	Coefficient	HAC standard error	p-value	N	Adjusted R^2
Conventional	Daily intercept	0.000191	0.000137	0.1643	3,032	0.9208
Conventional	Market beta	0.873249	0.015318	<0.0001	3,032	0.9208
Conventional	Market timing	-0.516315	0.897804	0.5652	3,032	0.9208
Conventional	Volatility timing	-0.012752	0.011740	0.2774	3,032	0.9208
Conventional	NAV-implied liquidity timing	0.022471	0.016110	0.1630	3,032	0.9208
Conventional	Co-skewness timing	0.022589	0.013918	0.1046	3,032	0.9208
Conventional	Co-kurtosis timing	-0.001948	0.010863	0.8577	3,032	0.9208
Islamic	Daily intercept	0.000173	0.000135	0.1982	3,037	0.8902
Islamic	Market beta	0.901890	0.016385	<0.0001	3,037	0.8902
Islamic	Market timing	-0.588709	0.853746	0.4905	3,037	0.8902
Islamic	Volatility timing	-0.001087	0.011368	0.9238	3,037	0.8902
Islamic	NAV-implied liquidity timing	0.022499	0.016845	0.1817	3,037	0.8902
Islamic	Co-skewness timing	0.007044	0.012325	0.5676	3,037	0.8902
Islamic	Co-kurtosis timing	-0.000004	0.011424	0.9997	3,037	0.8902

Source: Authors' estimates. HAC inference uses ten lags.

5.4 Regime evidence

The conditional coefficients are translated to make them more economically intuitive states using Table 6 and Figure 3. In low volatility states, conventional funds have a 2.89% annualized active return while in high volatility states, the annualized active return is -1.94%. They have a range of betas between 0.889 and 0.840, which is consistent with managing their exposure to volatility in a defensive way. The active performance is almost flat, and Islamic beta is relatively constant across volatility states. When the data is presented in the form of liquidity, a more discernible common pattern emerges. Conventional active return is 1.55% in high-liquidity states compared to -0.60% in low-liquidity states. The difference in returns is not significant between Islamic active return of 2.72% and -2.27% in high and low liquidity, respectively.

The condition of the joint stress is the most revealing. In the case where volatility is high and liquidity is low, annualized active return drops to -6.26% and -3.95% for conventional and Islamic funds, respectively. The volatility is increased to 21.62% and 23.63%, respectively. Neither system provides a complete insulation of the participants from a combined risk and valuation friction shock. This finding lends credence to the idea of evaluating managers simultaneously on liquidity and volatility: both measures acting, separately, overestimate the negative environment.

Table 6. Performance across volatility and liquidity regimes

System	Regime	Days	Market beta	Annualized active return (%)	Annualized volatility (%)
Conventional	Low volatility	1,516	0.889	2.89	16.22
Conventional	High volatility	1,516	0.840	-1.94	19.21
Conventional	Low liquidity	1,515	0.843	-0.60	18.90
Conventional	High liquidity	1,517	0.880	1.55	16.52
Conventional	Low vol / high liq	718	0.918	1.12	16.34
Conventional	High vol / low liq	717	0.833	-6.26	21.62
Islamic	Low volatility	1,518	0.899	-0.11	16.82
Islamic	High volatility	1,519	0.895	0.57	20.77
Islamic	Low liquidity	1,519	0.889	-2.27	20.34
Islamic	High liquidity	1,518	0.907	2.72	17.30
Islamic	Low vol / high liq	718	0.923	0.60	16.77
Islamic	High vol / low liq	719	0.896	-3.95	23.63

Source: Authors' calculations. Regimes use medians of lagged state variables.

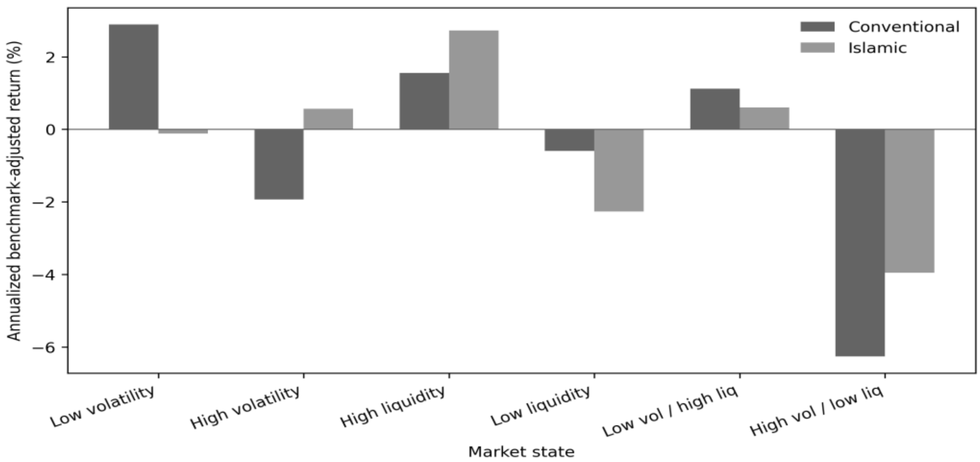

Figure 3. Annualized active returns across volatility-liquidity regimes. Source: Authors' calculations.

5.5 Machine-learning results

Table 7 reports the strict out-of-sample test. The historical-mean forecast has the lowest MAE and RMSE and, by construction, R^2_{OS} = 0. Every flexible model has a negative value, ranging from -0.0247 for Elastic Net to -0.1643 for Random Forest. Directional accuracy ranges from 40.6% to 45.8%, and prediction-return correlations are near zero or negative.

Table 7. Out-of-sample prediction of next-month active return, 2023-July 2026

Model	MAE	RMSE	Out-of-sample R^2	Directional accuracy	Prediction-return correlation
Historical mean	0.01968	0.02876	0.0000	0.4100	0.0000
Elastic Net	0.01999	0.02911	-0.0247	0.4191	0.0204
Extra Trees	0.02058	0.02993	-0.0830	0.4405	-0.0553
Ridge	0.02113	0.03024	-0.1056	0.4578	0.0491
Gradient Boosting	0.02139	0.03059	-0.1319	0.4486	-0.0288
Random Forest	0.02165	0.03103	-0.1643	0.4059	-0.1783

Source: Authors' calculations; 983 test fund-months. All predictors are lagged and the split is chronological.

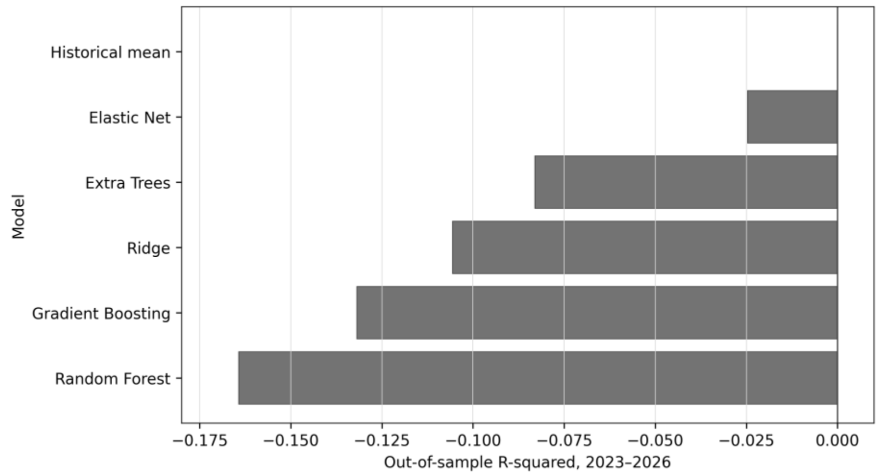

Figure 4. Out-of-sample R-squared by forecasting model relative to the historical-mean benchmark. Negative values indicate deterioration.

The forecasting failure is not merely a model-selection issue. Positive active-return observations decline from 52.9% in training to 41.0% in testing, indicating a shift in the conditional distribution. Tree ensembles, designed to capture nonlinear interactions, deteriorate more than penalized linear models. The evidence therefore does not support reliable monthly forecasting from the included states or fund identifiers. Econometric timing coefficients should be interpreted as historical conditional behaviour, not as an allocation rule.

5.6 Robustness

An aggregate coefficient robustness summary is provided in Table 8. In all three systems (42-day, 63-day, 126-day), the market timing indicators are negative. Volatility timing is still negative as is liquidity and co-skewness timing. Since the acute COVID-19 time period is excluded, the market-timing coefficients are rather more negative than the result. The most sensitive dimension is the co-kurtosis. The Islamic co-kurtosis coefficient, the result of the underlying regression, takes positive values and emerges statistically significant within the 126-

day time period, which is undesirable from the viewpoint of a tail-risk-avoider. This instability also indicates that rolling kurtosis is not a stable means of measuring skill, and should not be assumed to be an established standard of skill.

Table 8. Robustness of aggregate multidimensional timing coefficients

System	Specification	N	Market	Volatility	Liquidity	Co-skewness	Co-kurtosis
Conventional	42-day states	3,053	-0.5220	-0.0166	0.0210	0.0133	0.0069
Islamic	42-day states	3,058	-0.6151	-0.0013	0.0226	0.0055	0.0069
Conventional	63-day states	3,032	-0.5109	-0.0140	0.0216	0.0197	0.0001
Islamic	63-day states	3,037	-0.5959	-0.0012	0.0227	0.0062	0.0028
Conventional	126-day states	2,970	-0.5080	-0.0227	0.0242	0.0168	0.0115
Islamic	126-day states	2,975	-0.6684	-0.0039	0.0248	0.0165	0.0247
Conventional	63-day, excluding acute COVID	2,940	-0.9707	-0.0227	0.0118	0.0198	0.0066
Islamic	63-day, excluding acute COVID	2,945	-1.1611	-0.0041	0.0076	0.0041	0.0071

Source: Authors' estimates. Coefficients are shown to emphasize sign stability; full p-values are available from the estimation output.

6. Discussion, Contributions, and Implications

Market timing is not a single ability. Classical convexity tests provide no evidence of broad market-direction skill, and Islamic funds display more negative common-benchmark curvature. Yet conventional funds often reduce beta when volatility rises and increase exposure in favourable liquidity and skewness states. These patterns are more consistent with conditional risk management than directional speculation and align with Busse (1999), Ferson and Mo (2016), and Busse et al. (2024).

Liquidity conditions are economically important for both systems. Positive active returns in high-liquidity states and negative returns in low-liquidity states imply a state-dependent cost for beneficiaries. The NAV proxy has the potential to reflect both market liquidity and operational repricing, making it a governance issue rather than a measure of bid ask spreads. There is a need to incentivise regulators to require disclosure of valuation timing, stale price, cash buffers, turnover and liquidity stress tests.

Co-skewness is a more useful indicator than co-kurtosis. The difference between conventional and most Islamic funds is statistically significant and the exposures of the funds are higher when the lagged market asymmetry is more favourable. But this impact is null in portfolio of aggregates. The timing of kurtosis is weaker and less stable, as the timing of unsigned kurtosis is difficult to estimate precisely within short time frames.

The comparison between conventional and Islamic is not a comparison that should be narrowed down to a better versus worse ranking. While Islamic funds are less stable, they do not have any notably lower returns, return-to-volatility ratios, or worse drawdowns. The less desirable common-benchmark curvature might be due to structural screening, not to bad judgment. There should therefore be a common opportunity-cost-based benchmark as well as a Shariah-consistent benchmark that can be reported in the fair evaluation.

The article has four contributions. It offers day-by-day insights on market timing, volatility timing, liquidity timing, co-skewness timing, and co-kurtosis timing both for the conventional and Islamic pension funds in Pakistan; it introduces a transparent NAV-implied liquidity and valuation-friction state under settings where holdings data is unavailable; it integrates HAC inference with false-discovery control; and imposes a rigorous out-of-sample machine-learning test on the state variables. The failure of forecasting is informative in that it distinguishes between conditional association and stable predictive value.

The results help to create a multi-dimensional scorecard for pension trustees, which includes benchmark beta, downside performance, stress-regime active returns, liquidity exposure and multiple-testing adjusted timing evidence. The regulators would welcome uniformity of NAV definition, validated expense ratios, official total-return archives, portfolio-liquidity metrics and holdings disclosure. The actual difficulty for managers is to understand how to manage the exposure in a market when volatility and liquidity are decreasing simultaneously.

The study has some limitations. The primary models provide information on total rather than excess returns; a daily risk-free series should be included. The benchmark is not a system specific, but common one. The liquidity state is a NAV implied state, and is not able to distinguish trading costs from stale valuation. There is no information available about the size of the funds, the amount of funds flowing, what funds the fund holds, how long the managers have been in place, or fund transaction costs. The panel could be survivorship and backfill effects and equal weight portfolios are not assets under management. The post-2022 machine-learning test is not as long and its macroeconomic regime varies. These constraints serve to focus the conclusions so that they are intentionally more constrained—instead of private information or causal managerial talent, conditional return behavior.

7. Conclusion and Future Research

Within this larger source panel of 316,448 daily observations, this study assesses 24 pension equity funds out of which 11 is conventional and 13 is Islamic pension fund. The benchmark-matched analysis includes 61,520 fund-day data observations between January 2014 and July 2026. Most of the classical Treynor–Mazuy, Henriksson–Merton and integrated timing coefficients are negative and fail the false-discovery correction. Conditional tests produce more nuanced results: conventional funds exhibit positive average volatility, liquidity and co-skewness timing while Islamic funds exhibit positive median timing in several dimensions, but with less cross-sectional uniformity. Islamic funds are more volatile, but general return and general drawdown do not create system dominance. Both systems fail if the volatility is high, and the liquidity is low.

The out of sample exercise is equally important for a qualification. All flexible models underperformed the historical-mean benchmark, which means that conditional relations seen in-sample should not be used as readily-exploitable forecasts. There is evidence that the exposure behaviour is regime-sensitive and even heterogeneous, rather than a simple trading rule.

Future studies should include turnover, fund size, flows, holdings, change in manager, risk-free data from State Bank of Pakistan, Amihud illiquidity and spreads, and compare the KSE-100 and KMI-30 Total Return indices, with expenses being validated. Interest-rate timing models are needed for debt and money-market sub-funds, and gold and exchange-rate factors for commodity sub-funds. Structural screening may be further differentiated from active management using time-varying-parameter, hidden-Markov modelling, and dynamic-correlation or economically constrained machine-learning models. Survivorship and stale pricing issues should be addressed with the introduction of discontinued funds, as well as documented NAV-pricing procedures.

Declarations

Acknowledgements. The author acknowledges the pension-fund NAV data and PSX total-return data used in this study.
Data availability. Fund-level source workbooks were supplied for this research, and KSE-1000 Total Return Index data were obtained from Pakistan Stock Exchange historical files. Processed data and estimation outputs may be available on reasonable request, subject to data-owner restrictions.
Funding. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.
Conflict of interest. The author declares no conflict of interest.
Ethical approval. The study uses institutional financial data and does not involve human participants or personal data.

References

Alqahtani, M., Henry, D., & Bhatti, M. I. (2025). The performance Islamic versus conventional mutual funds during extreme event: New evidence from Saudi Arabia. *Journal of Islamic Accounting and Business Research*. Advance online publication. https://doi.org/10.1108/JIABR-06-2025-0325

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31-56.

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. https://doi.org/10.1111/j.1540-6261.2009.01527.x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B, 57*(1), 289-300.

Bollen, N. P. B., & Busse, J. A. (2001). On the timing ability of mutual fund managers. *Journal of Finance, 56*(3), 1075-1094. https://doi.org/10.1111/0022-1082.00356

Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5-32.

Busse, J. A. (1999). Volatility timing in mutual funds: Evidence from daily returns. *Review of Financial Studies, 12*(5), 1009-1041. https://doi.org/10.1093/rfs/12.5.1009

Busse, J. A., Ding, J., Jiang, L., & Tang, Y. (2023). Artificial market timing in mutual funds. *Journal of Financial and Quantitative Analysis, 58*(8), 3450-3481. https://doi.org/10.1017/S0022109022001363

Busse, J. A., Ding, J., Jiang, L., & Wu, K. (2024). Dynamic market timing in mutual funds. *Management Science, 70*(6), 3470-3492. https://doi.org/10.1287/mnsc.2023.4857

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out-of-sample: Can anything beat the historical average? *Review of Financial Studies, 21*(4), 1509-1531.

Cao, C., Simin, T., & Wang, Y. (2013). Do mutual fund managers time market liquidity? *Journal of Financial Markets, 16*(2), 279-307. https://doi.org/10.1016/j.finmar.2012.10.004

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*(1), 57-82.

Chang, E. C., & Lewellen, W. G. (1984). Market timing and mutual fund investment performance. *Journal of Business, 57*(1), 57-72.

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794).

Chu, T., Weng, C., Tsai, C.-F., & Li, C.-T. (2022). Predicting mutual funds' performance using deep learning and ensemble techniques. *arXiv*. https://doi.org/10.48550/arXiv.2209.09649

Climent, F., Momparler, A., & Carmona, P. (2025). Predicting mutual fund performance with machine learning: Are ESG pillar scores relevant predictors of fund return? *Borsa Istanbul Review, 25*(6), 1403-1419. https://doi.org/10.1016/j.bir.2025.07.019

Delle Foglie, A., & Panetta, I. C. (2020). Islamic stock market versus conventional: Are Islamic investing a safe haven for investors? A systematic literature review. *Pacific-Basin Finance Journal, 64*, 101435. https://doi.org/10.1016/j.pacfin.2020.101435

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics, 13*(3), 253-263.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics, 20*(3), 339-350.

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. https://doi.org/10.1111/j.1540-6261.2010.01598.x

Fang, H., & Lai, T.-Y. (1997). Co-kurtosis and capital asset pricing. *Financial Review, 32*(2), 293-307.

Ferson, W. E., & Mo, H. (2016). Performance measurement with selectivity, market and volatility timing. *Journal of Financial Economics, 121*(1), 93-110. https://doi.org/10.1016/j.jfineco.2016.02.012

Ferson, W. E., & Schadt, R. W. (1996). Measuring fund strategy and performance in changing economic conditions. *Journal of Finance, 51*(2), 425-461. https://doi.org/10.1111/j.1540-6261.1996.tb02690.x

Foran, J., & O'Sullivan, N. (2017). Mutual fund skill in timing market volatility and liquidity. *International Journal of Finance & Economics, 22*(4), 257-273. https://doi.org/10.1002/ijfe.1580

Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *Review of Financial Studies, 33*(5), 2223-2273.

Harvey, C. R., & Siddique, A. (2000). Conditional skewness in asset pricing tests. *Journal of Finance, 55*(3), 1263-1295.

Henriksson, R. D., & Merton, R. C. (1981). On market timing and investment performance. II. Statistical procedures for evaluating forecasting skills. *Journal of Business, 54*(4), 513-533. https://doi.org/10.1086/296144

Jagannathan, R., & Korajczyk, R. A. (1986). Assessing the market timing performance of managed portfolios. *Journal of Business, 59*(2), 217-235.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *Journal of Finance, 23*(2), 389-416.

Jiang, W. (2003). A nonparametric test of market timing. *Journal of Empirical Finance, 10*(4), 399-425.

Kacperczyk, M., Van Nieuwerburgh, S., & Veldkamp, L. (2014). Time-varying fund manager skill. *Journal of Finance, 69*(4), 1455-1484.

Khaleel, F., Ihsan, H., & Rashid, A. (2026). Do ownership structure and governance matter in asset allocation decisions of Islamic and conventional mutual funds? Empirical evidence from Pakistan. *Borsa Istanbul Review, 26*(1), 100745. https://doi.org/10.1016/j.bir.2025.10.018

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. https://doi.org/10.1111/j.1540-6261.2006.01015.x

Kraus, A., & Litzenberger, R. H. (1976). Skewness preference and the valuation of risk assets. *Journal of Finance, 31*(4), 1085-1100.

Lesmond, D. A., Ogden, J. P., & Trzcinka, C. A. (1999). A new estimate of transaction costs. *Review of Financial Studies, 12*(5), 1113-1141.

Mohammad, N., & Ashraf, D. (2015). The market timing ability and return performance of Islamic equities: An empirical study. *Pacific-Basin Finance Journal, 34*, 169-183. https://doi.org/10.1016/j.pacfin.2015.07.001

Momparler, A., Carmona, P., & Climent, F. (2025). Mutual fund performance prediction using machine learning. *Computational Economics, 65*(3), 1617-1642. https://doi.org/10.1007/s10614-024-10618-0

Moreira, A., & Muir, T. (2017). Volatility-managed portfolios. *Journal of Finance, 72*(4), 1611-1644.

MUFAP. (2026). *Voluntary pension schemes*. Mutual Funds Association of Pakistan.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*(3), 703-708.

Pakistan Stock Exchange. (2026). *Historical KSE-30 Total Return Index data and index methodology*. PSX Data Portal.

Pastor, L., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy, 111*(3), 642-685.

SECP. (2023). *Khyber Pakhtunkhwa transitions newly hired civil servants to the Voluntary Pension System*. Securities and Exchange Commission of Pakistan.

SECP. (2024a). *SECP amends VPS framework to allow fully funded defined contribution pension schemes* [Press release, 1 March 2024]. Securities and Exchange Commission of Pakistan.

SECP. (2024b). *Workshop on the new regulatory framework of Voluntary Pension Schemes* [Press release, 4 March 2024]. Securities and Exchange Commission of Pakistan.

Treynor, J. L., & Mazuy, K. K. (1966). Can mutual funds outguess the market? *Harvard Business Review, 44*, 131-136.

Wattanatorn, W., & Padungsaksawasdi, C. (2021). Cokurtosis and the ability of mutual fund managers. *Finance Research Letters, 40*, 101777. https://doi.org/10.1016/j.frl.2020.101777

Yin, C., O'Sullivan, N., & Sherman, M. (2024). Mutual fund liquidity timing and the choice of liquidity proxy. *North American Journal of Economics and Finance, 74*, 102201. https://doi.org/10.1016/j.najef.2024.102201