

State-Dependent Asset Pricing and Machine-Learning Return Forecasting: Evidence from the Pakistan Stock Exchange



Kainat Arif ^a Muhammad Naveed Jan ^b Usman Ayub ^c

Abstract: *This study asks whether asset-pricing relationships and return predictability in an emerging market vary across market conditions, and whether machine-learning methods capture this state dependence more flexibly than linear approaches. Using 18 Pakistan Stock Exchange (PSX) portfolio-return series (February 2010-December 2025, $N=191$), we estimate Fama-French five-factor (FF5) and Arbitrage Pricing Theory (APT)-motivated macroeconomic specifications with ordinary least squares (OLS), an artificial neural network (ANN) and Random Forest (RF), separately for the full sample and for rolling-window-classified Bullish ($n=26$), Bearish ($n=30$) and Neutral ($n=124$) market conditions. The FF5/OLS specification has substantial in-sample explanatory power (average $R^2=0.397$); the macroeconomic/OLS specification is statistically significant for several variables but has much lower explanatory power (average $R^2=0.061$). On a 70/30 chronological holdout, Random Forest extracts considerable nonlinear predictive structure from the FF5 factors (average out-of-sample $R^2=0.645$; mean directional accuracy=80.3%), while the ANN is more sensitive to the limited sample ($R^2=0.055$). Macroeconomic machine-learning models do not generalise out of sample (negative average R^2). Factor and feature importance shift materially across market conditions. These patterns are consistent with state-dependent asset pricing and the Adaptive Market Hypothesis, but the evidence is conditional on the sample, predictor set and model specification, and does not establish universal machine-learning superiority or investment profitability.*

Keywords: Asset Pricing, Machine Learning, Random Forest, Artificial Neural Network, Adaptive Market Hypothesis, Fama-French Five-Factor Model, Arbitrage Pricing Theory, Emerging Markets, Pakistan Stock Exchange, Return Predictability

JEL Classification: G12; G14; G17; C45; C53

Introduction

It is important to know whether the returns on stocks are predictable and the type of risk factors responsible for it. The capital asset pricing model and the multi-factor extensions, which include the Arbitrage Pricing Theory (APT; Ross, 1976) and the Fama and French Five-Factor Model (FF5; Fama & French, 2017), have been the main linear models used to explain the relationship between the expected return and systematic risk based on firm characteristics or macroeconomic state variables. All the above models are based on the assumption that there is a linear, stable and time-invariant relationship between factor returns. There has been a lot of work done lately to challenge this assumption and instead show that the predictability of the returns, factor premium and model fit change depending on the existing market environment, as per the Adaptive Market Hypothesis (AMH; Lo, 2004).

State dependency in any return generating process means that an aggregate linear regression would only approximate an average relationship which would never fit to any single market state. Machine learning

^a Research Scholar, Department of Management Sciences, COMSATS University Islamabad, Abbottabad Campus, Khyber Pakhtunkhwa, Pakistan.

^b Assistant Professor, Department of Management Sciences, COMSATS University Islamabad, Abbottabad Campus, Khyber Pakhtunkhwa, Pakistan.

^c Professor, Department of Economics, COMSATS University Islamabad, Abbottabad Campus, Khyber Pakhtunkhwa, Pakistan.

techniques, especially artificial neural networks (ANN) and Random Forest (RF), are natural complements to this question since they allow capturing nonlinear effects, threshold effects, and interaction between predictors, without the researcher pre-specifying functional form (Gu et al., 2020; Fallahgoul et al., 2024; Kelly et al., 2024). Machine learning is not the theory of asset pricing but rather an estimation technique that can help in testing theoretical predictions about the relationship between economic variables and returns that cannot be captured by a linear regression model, and whether this relationship varies between market states.

An emerging market is an appropriate place to study such issues. The Pakistan Stock Exchange (PSX) exhibits ongoing pressure due to inflation, exchange rate uncertainty, energy supply restrictions, and times of heightened political risk alongside characteristics like thinner trading, fewer stocks on the listing, and incomplete disclosures that are likely to result in non-linearities and regime-dependent return behavior that differs from what we see in developed and large emerging markets. Most asset pricing studies conducted at the PSX have used different risk factors in separate studies, mixed data across market conditions, and relied on old-fashioned machine learning models instead of recent ensemble and neural network techniques (Hanauer & Kalsbach, 2023; Raza et al., 2024; Ullah et al., 2025).

This paper fills three connected holes in the literature. First, it tests the FF5 factors and APT-inspired seven macroeconomic factors as two distinct sets of theoretically based predictor variables by employing the same set of 18 portfolios from the PSX market and the same three methodologies of estimation (OLS, ANN, RF) so that their explanatory and predictive value could be evaluated in a systematic fashion rather than mixed up in one regression specification. Second, it contrasts OLS with ANN and Random Forest models in terms of out-of-sample performance using a chronological out-of-sample test design, and it makes sure not to confuse the explanatory power of in-sample regression with the predictive ability of out-of-sample forecasting. Third, it classifies the sample into Bullish, Bearish and Neutral market conditions using a transparent, rolling-window rule and estimates each model separately within each condition, allowing a direct empirical assessment of whether factor relevance, model fit and feature importance are state dependent, consistent with the AMH.

The central empirical question is not whether one algorithm is universally best. It is whether return predictability in the PSX is predictor-set dependent, model dependent and market-condition dependent — that is, whether the answer to “what explains PSX returns” and “how well can PSX returns be forecast” changes depending on which risk factors, which estimation method and which market regime are considered. Anticipating the results, FF5 factors carry substantial explanatory relevance, macroeconomic variables are statistically related to returns but explain much less of their variation, Random Forest extracts material nonlinear structure from the FF5 factors under the selected out-of-sample design, ANN is comparatively more sensitive to the limited monthly sample, macroeconomic machine-learning models fail to generalise out of sample, and factor and feature importance shift materially across Bullish, Bearish and Neutral conditions.

The study contributes to the literature in four ways. Empirically, it provides portfolio-level, PSX-specific evidence spanning a 15-year period that includes the post-financial-crisis recovery, sustained growth, the COVID-19 shock and a subsequent inflationary and currency-depreciation episode. Methodologically, it evaluates OLS, ANN and RF under a common chronological holdout, and is explicit about the difference between in-sample and out-of-sample evaluation rather than treating the two as directly comparable. Conceptually, it integrates the APT, FF5 and AMH within a single empirical architecture, using machine learning as the estimation tool that allows the state dependence implied by the AMH to be examined empirically. Finally, it documents a comprehensive negative result — the failure of macroeconomic machine-learning models to generalise out of sample — that is itself informative about the structural instability of macro-return relationships in a volatile emerging market.

Literature Review and Theoretical Development

Traditional Asset Pricing: CAPM, APT and Fama-French Models

The capital asset pricing model prices assets through covariance with the market portfolio; Ross's (1976) APT relaxes this to allow multiple, unspecified systematic risk factors, motivating the use of macroeconomic state variables — inflation, money supply, exchange rates, industrial production, oil prices and policy rates — as candidate priced risks (Chen & Jordan, 1993; Priestley, 1996; Connor & Korajczyk, 1995). Fama and French's (2017) international five-factor extension adds size, value, profitability and investment dimensions to the market factor, and has been tested across developed and emerging markets with mixed support for individual factors (Chai et al., 2016; Sha & Gao, 2019). In Pakistan, Haqqani and Aleem (2020) extend the FF5 framework with a liquidity factor, and Khan, Khan, and Saeed (2023) propose a human-capital-based four-factor model, both finding that firm-specific characteristics carry independent pricing content in the PSX beyond standard factors. Macroeconomic and firm-specific factors are best viewed as complementary rather than competing sources of pricing information (Heet et al., 2015a, 2015b), which motivates evaluating them as separate predictor sets in this study rather than combining them within a single specification.

Machine Learning in Asset Pricing

Machine-learning methods relax the linear, additive structure of factor models. Gu et al. (2020) show that tree ensembles and neural networks achieve materially higher out-of-sample explanatory power than OLS for U.S. equity returns by capturing nonlinear characteristic interactions; Kelly, Malamud, and Zhou (2024) show that increasing model complexity can improve return prediction even in relatively short samples; Chen et al. (2024) embed no-arbitrage restrictions within a deep-learning asset-pricing model; and Fallahgoul, Frantiano, and Lin (2024) develop significance tests for neural-network asset-pricing models, addressing the inferential difficulty of standard machine-learning output. Bagnara et al. (2024) critical review concludes that these methods can usefully complement, but do not replace, theory-motivated factor models — a conclusion this study's design is built to test directly rather than assume.

Artificial Neural Networks in Financial Prediction

ANNs have a long history in return and price prediction (Armano et al., 2005; Cao et al., 2011; Carvalhal & Ribeiro, 2008; Idowu et al., 2012), including early Pakistan-specific applications (Fatima & Hussain, 2008; Jan, 2019) that predate the deep-learning architectures and inferential methods now used internationally. More recent work continues to document that ANN performance is sensitive to sample size, architecture and initialisation, particularly in monthly-frequency financial data where the number of training observations is small relative to the number of network parameters (Kazak et al., 2025; Liang et al., 2024). This sensitivity is directly relevant to the present study's limited monthly sample and market-condition subsamples.

Random Forest and Ensemble Learning

Random Forest (Breiman, 2001) aggregates many decision trees to reduce variance and capture nonlinear interactions without an explicit functional form, and has performed well relative to linear and neural-network benchmarks in medium-sized financial datasets (Yao et al., 2022; Fang et al., 2021; Cheng et al., 2025). Ensemble tree-based variable-importance measures have also been used to select parsimonious, predictively accurate factor specifications in Pakistan (Ullah et al., 2025). Feature importance in RF reflects the average reduction in prediction error attributable to a variable across trees; it is a measure of predictive association within the fitted model, not a causal or structural parameter, and is interpreted accordingly throughout this study.

State-Dependent and Regime-Dependent Asset Pricing

A separate literature documents that risk premia, factor loadings and predictability vary with the state of the market. Nyberg (2013) and Zegadło (2022) model bull and bear regimes directly; Zaremba and Szczygielski (2019) show that limits to arbitrage and sentiment affect factor returns state-dependently in international

government bond markets; and Bradrania and Pirayesh Neghab (2022) find that ANN architectures conditioned on market-state variables improve both return prediction and the economic value of portfolio construction relative to unconditional models. These findings motivate estimating OLS, ANN, and RF separately within Bullish, Bearish, and Neutral PSX conditions rather than pooling across states, and they anticipate the AMH-consistent finding that factor and feature importance shift across conditions.

Emerging-Market and Pakistan Stock Exchange Evidence

Hanauer and Kalsbach (2023) find that machine-learning return prediction performs well across emerging markets as a group, without establishing PSX-specific applicability. Munir et al. (2023) and Raza et al. (2024) apply machine-learning methods to PSX price and return prediction using technical indicators, and Ullah et al. (2025) compare machine-learning-selected factor specifications against conventional criteria; none of these studies integrates FF5 and macroeconomic factor modelling, explicit Bullish/Bearish/Neutral segmentation, and chronological out-of-sample machine-learning evaluation within a single, common portfolio universe.

Research Gap

Taken together, the literature suggests three propositions this study tests directly. First, FF5 firm-level factors are likely to explain PSX portfolio returns more fully than APT-motivated macroeconomic variables, consistent with international evidence that macroeconomic factor models generally underperform firm-characteristic models in explanatory power (He et al., 2015a). Second, nonlinear machine-learning methods, and Random Forest in particular, are plausible candidates for capturing predictive structure in the FF5 factors that a linear specification leaves unexploited, though their relative performance against ANN is expected to depend on sample size given the short monthly series available. Third, the relevance of factors, the fit of the model and its predictive performance are expected to vary significantly between Bullish, Bearish and Neutral regimes, in line with AMH as well as international regime-dependent literature. These hypotheses, along with the clear-cut delineation between in-sample and out-of-sample performance, represent the empirical structure outlined below.

Methodology

Research Design

This is a quantitative, empirical study that incorporates in-sample factor model (OLS) estimation with out-of-sample machine learning prediction (ANN, Random Forest), done separately for two theoretical predictors (FF5 factors; APT-based macroeconomic variables) and for the full sample and three market scenarios (Bullish, Bearish, Neutral). OLS is used for coefficient estimation, statistical inference and in-sample explanatory analysis; ANN and Random Forest are evaluated on a held-out chronological test sample. Because in-sample and out-of-sample R^2 are computed on different observations and measure different properties, they are reported and interpreted separately rather than ranked on a single scale.

Market, Sample and Data Sources

The dataset consists of 18 monthly return series of the PSX portfolios (P1_1-P1_6, P2_1-P2_6, P3_1-P3_6), which consist of nonfinancial companies of the KSE-100 index, provided in ready-made form. The portfolios have been divided into 3 categories on the basis of size (P1, P2, P3), while each category is further subdivided into six sub-groups on the basis of risk and return. The information about the basis for dividing portfolios into 3 categories is provided in the underlying project document, while the basis for dividing each category into 6 sub-groups is not clearly mentioned. Monthly portfolio-return and FF5 factor data are drawn from standard financial databases; macroeconomic data are sourced from the State Bank of Pakistan (policy rate, M2), the Pakistan Bureau of Statistics (CPI, industrial production index), and the World Bank/IMF (annual real GDP growth, linearly interpolated to monthly frequency to align with the return series). The analysis period is January 2010-December 2025 (192 raw monthly observations), chosen to span multiple

market cycles: post-financial-crisis recovery (2010-2012), sustained growth (2013-2017), rising political and macroeconomic uncertainty (2018-2019), the COVID-19 shock (2020-2021), the subsequent inflationary and currency-depreciation episode (2022-2023), and partial stabilization (2024-2025). No observations were excluded, winsorized, or adjusted for the COVID-19 period; extreme-event months were retained and classified using the same rolling-window rule as all other months, consistent with the AMH view that disruptive episodes are informative rather than anomalous.

First-differencing the flow-based macroeconomic variables (oil price, M2, exchange rate, industrial production index) removes one observation, leaving 191 monthly observations (February 2010-December 2025) for the full-sample OLS and machine-learning models. Because the market-condition classification requires a 12-month rolling window, the first 12 months of the sample cannot be classified, leaving 180 monthly observations (January 2011-December 2025) for the market-condition-specific models, comprising 26 Bullish (14.4%), 30 Bearish (16.7%) and 124 Neutral (68.9%) months. Sampling is exhaustive (census) over the specified period rather than a drawn sub-sample.

Variables

Table 1 defines the dependent variable, the five FF5 factors and the seven macroeconomic variables used as two separate predictor sets.

Table 1

Variable Definitions, Measurement and Data Sources

Variable	Definition / measurement	Source
Portfolio return (dependent)	$R_{p,t} = (P_{p,t} - P_{p,t-1}) / P_{p,t-1} \times 100$, monthly %, 18 series	PSX; own calculation
MKT	Market excess return: KSE-100 monthly return minus risk-free rate (T-bill/SBP repo)	PSX; SBP
SMB	Size factor: small-cap minus large-cap portfolio return	Constructed (Fama & French, 2017)
HML	Value factor: high book-to-market minus low book-to-market return	Constructed
RMW	Profitability factor: robust minus weak operating-profitability return	Constructed
CMA	Investment factor: conservative minus aggressive asset-growth return	Constructed
Oil_ret	Monthly % change, international oil price	World Bank / IMF
M2_ret	Monthly % change, broad money supply	State Bank of Pakistan
GDP	Real GDP growth, annual, linearly interpolated to monthly frequency	World Bank / IMF
ExRate_ret	Monthly % change, PKR/USD exchange rate	State Bank of Pakistan
PolicyRate	SBP target policy rate, monthly	State Bank of Pakistan
CPI	Monthly % change, Consumer Price Index	Pakistan Bureau of Statistics
IPI_ret	Monthly % change, Industrial Production Index	Pakistan Bureau of Statistics

Note: Macroeconomic flow variables are transformed to percentage changes for stationarity, which removes one observation from the full estimation sample. GDP interpolation smooths short-run variation and does not represent independently observed monthly GDP; this is a documented data limitation

Market-Condition Classification

Market conditions are classified using a rolling-window rule applied to the monthly market return (MKT). For each month t (from month 13 onward), a 12-month rolling mean and standard deviation are computed:

$$\mu_t = (1/12) \sum \text{MKT}_{t-i}, \quad \sigma_t = \sqrt{[(1/12) \sum (\text{MKT}_{t-i} - \mu_t)^2]}, \quad i = 0, \dots, 11$$

with the classification rule: Bullish if $MKT_t > \mu_t + \sigma_t$; Bearish if $MKT_t < \mu_t - \sigma_t$; Neutral otherwise. This is an operational, transparent classification rule rather than a structural Markov regime-switching model: it does not estimate unobserved states or transition probabilities, and it does not smooth gradual transitions between conditions. A methodologically important feature of this rule, which we make explicit here, is that the summation index runs from $i = 0$, so month t 's own return enters the rolling mean and standard deviation used to classify month t . The classification is therefore contemporaneous rather than a lagged, ex-ante signal: it should be understood as a state-conditioned asset-pricing analysis — partitioning the sample by realised market conditions to study state-dependent pricing and predictive relationships — rather than as an ex-ante market-timing or regime-forecasting model that an investor could have implemented in real time without knowledge of the current month's return. This distinction does not affect the paper's central explanatory and predictive comparisons, which are conducted within each classified condition, but it does mean the market-condition labels themselves should not be interpreted as forecastable in advance without further, lagged-classification analysis.

No attempt was made to balance the three conditions' observation counts, since doing so would require deleting, duplicating or resampling actual time-series observations and would risk introducing look-ahead or selection bias; the resulting class imbalance (26 Bullish, 30 Bearish, 124 Neutral) is a direct, unadjusted consequence of the classification rule applied to PSX return history.

OLS Specifications

For each of the 18 portfolios, two OLS specifications are estimated with heteroskedasticity-consistent standard errors, first with the FF5 factors:

$$R_{p,t} = \alpha_p + \beta_1 MKT_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 RMW_t + \beta_5 CMA_t + \varepsilon_{p,t}$$

and separately with the seven macroeconomic variables:

$$R_{p,t} = \alpha_p + \gamma_1 Oil_t + \gamma_2 M2_t + \gamma_3 GDP_t + \gamma_4 ExRate_t + \gamma_5 PolicyRate_t + \gamma_6 CPI_t + \gamma_7 IPI_t + \varepsilon_{p,t}$$

Each specification is estimated separately for the full sample and for the Bullish and Bearish sub-samples (Neutral is not separately modelled given the study's focus on extreme conditions). α_p and the factor loadings are interpreted as coefficient estimates within the specified model; joint significance is assessed with the F-statistic, and non-normality in the dependent variable motivates the use of heteroskedasticity-robust inference throughout.

Artificial Neural Network

The ANN is a multilayer perceptron with two hidden layers (64 and 32 units), ReLU activation, the Adam optimiser, mean-squared-error loss, early stopping and a maximum of 1,000 training epochs. Predictors are standardised using training-set statistics only. This architecture was pre-specified rather than exhaustively optimised through a time-series-aware hyperparameter search, and results should be described as reflecting

Random Forest

The Random Forest uses 500 trees with maximum depth 5, splitting on mean-squared-error reduction, and aggregates predictions by averaging across trees. As with the ANN, this configuration is a pre-specified benchmark rather than the product of exhaustive hyperparameter optimisation. Feature importance is computed as the average normalised impurity reduction attributable to each predictor across all trees and is reported as a measure of predictive association within the fitted ensemble, not as a causal or structural estimate of a variable's contribution to returns.

Train/Test Design and Forecast Evaluation

ANN and Random Forest are evaluated using a 70/30 chronological holdout (training $N=133$, test $N=58$ in the full sample; smaller, proportionally split sub-samples within Bullish and Bearish conditions), with the

earliest 70% of observations used for training and the latest 30% for testing, preserving temporal ordering. This single holdout does not constitute exhaustive rolling-origin or expanding-window validation, and the reported machine-learning performance is therefore conditional on this specific forecast origin, sample period, predictor set and fixed hyperparameters; a rolling-window robustness check for the best-performing portfolio partially addresses, but does not fully resolve, this limitation. Forecast accuracy is assessed with out-of-sample R^2 , root-mean-squared error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and mean directional accuracy (MDA, the percentage of test-period months in which the model correctly signs the return). MAPE is reported because it appears in the underlying analysis, but because monthly portfolio returns are frequently close to zero, MAPE is unstable and is given less interpretative weight than RMSE, MAE, out-of-sample R^2 and MDA.

A Note on the OLS-Machine-Learning Comparison

OLS and the machine-learning models serve different analytical purposes in this study. OLS is used for coefficient estimation, statistical inference and in-sample explanatory analysis; ANN and Random Forest are evaluated using out-of-sample forecasting metrics on the chronological test sample. Because in-sample and out-of-sample R^2 are computed on different observations and measure different properties of a model, they are not treated as directly comparable, and no ranking such as “Random Forest is better than OLS because its R^2 is higher” is drawn from this design. A fair, apples-to-apples comparison would require OLS predictions evaluated on the identical 30% chronological test sample used for ANN and Random Forest; this was not available in the underlying project materials reviewed for this article. We therefore report OLS as an explanatory/inference benchmark and ANN/Random Forest as forecasting models, and we identify a same-test-period OLS benchmark as a critical pre-submission robustness check.

Results

Descriptive Statistics

The 18 portfolios differ substantially in both return and risk (Table 2; Figure 2). Mean monthly returns range from -0.01% (P1_2) to 2.65% (P3_1), with P3-group portfolios earning consistently higher mean returns (0.84% - 2.65%) than P1 (-0.01% - 1.41%) and P2 (0.23% - 0.83%) portfolios. Standard deviations increase from 6.5% - 10.0% (P1) and 6.5% - 8.3% (P2) to 10.6% - 13.9% (P3), so increases in returns in the P3 group come at the expense of increasing standard deviation instead of a risk-free adjustment. Fifteen out of 18 portfolio groups have positive skewness and all have positive excess kurtosis, representing leptokurtic and positively skewed distributions seen in emerging equity market returns; the Jarque-Bera normality test fails in 16 out of 18 portfolios at the 5% level (except for P1_5 and P2_5 portfolios), justifying heteroskedasticity robust inference across the board. Maximum and minimum returns for the entire sample period occurred in March-April 2020 during the COVID-19 disruption and recovery.

Table 2

Descriptive Statistics, 18 PSX Portfolio Return Series (Monthly %, N=191, Feb 2010-Dec 2025)

Portfolio	Mean	Std Dev	Skewness	Ex. Kurtosis	JB p-value
P1_1	1.41	6.81	1.37	3.46	0.000
P1_5	1.28	7.84	0.23	0.34	0.263
P2_2	0.53	6.52	0.52	0.72	0.002
P2_5	0.83	7.95	0.35	0.01	0.136
P3_1	2.65	10.65	0.97	1.47	0.000
P3_6	2.13	13.87	0.79	0.22	0.000

Note: Ex. Kurt. = excess kurtosis (kurtosis – 3). JB = Jarque-Bera normality test p-value. Full 18-portfolio table available from the authors.

Market-Condition Classification

In the case of rolling-window method, out of 180 usable months, 26 (14.4%), 30 (16.7%) and 124 (68.9%) months are classified as Bullish, Bearish and Neutral respectively; the inequality results not from the sampling

but from the classification itself because most monthly returns fall within one standard deviation range from the trailing 12-month mean. The average difference in portfolio return between Bullish and Bearish months is economically significant: on average, returns are 7.83% and –5.94% in Bullish and Bearish months, respectively, giving us a gap of 13.77%.

FF5 Pricing Results

For all 18 portfolios, the FF5 regression is significant at the 1% level (F-test). MKT is significant at the 1% level in all portfolios, and the coefficients increase monotonically from an average of 0.46 (P1) to 0.78 (P2) and 1.40 (P3), as does the model's explanatory power, increasing from an average of 0.145 (P1) to 0.375 (P2) to 0.607 (P3), and an average for the whole sample of 0.397. The SMB factor coefficient is negative in all portfolios and significant in 15 out of 18 portfolios, with an average of –0.45, indicative of the PSX portfolio formation approach that favors large and liquid stocks. HML is negative in 14 of 18 portfolios (significant in 7, with coefficients between –0.16 and –0.41), indicating a growth rather than value tilt over this sample. RMW and CMA are significant mainly in the P3-group portfolios, suggesting that profitability and investment dimensions are more relevant pricing factors for larger, more institutionally scrutinised portfolios.

Table 3

Full-Sample FF5 OLS Results, 18 Portfolios (Selected)

Portfolio	MKT	SMB	HML	RMW	CMA	R ²	F-stat
P1_1	0.46**	–0.35	–0.12	0.209**	0.207**	0.118	4.94***
P2_1	0.78***	–0.39***	–0.21**	0.059	–0.225**	0.427	27.61***
P3_1	1.37***	–0.79***	–0.22**	0.091	0.257**	0.691	82.62***
P3_6	1.65***	–0.66***	–0.32***	0.119	0.461***	0.600	55.57***

Note: *** p<0.01, ** p<0.05, * p<0.10. Coefficients reported in decimal form. Full 18-portfolio coefficient matrix available from the authors.

Bullish- and Bearish-conditioned FF5 regressions show systematic differences from the full-sample results. In the Bullish sub-sample (N=26), the average MKT loading falls to 0.717 while CMA rises sharply (0.197 versus 0.072 full-sample) and HML becomes more negative (–0.253), consistent with a stronger growth tilt and a more prominent investment-conservatism premium during expansions. In the Bearish sub-sample (N=30), the average MKT loading rises to 0.876 — the highest of the three conditions — consistent with the well-documented flight-to-beta effect in which asset co-movement with the market intensifies during downturns. Average R² is higher in the Bullish condition (0.453) than the full sample, but adjusted R² falls sharply (0.316) reflecting the small sub-sample size; joint model significance holds in only 61% of Bullish models and 44% of Bearish models, again reflecting limited condition-specific sample sizes, although MKT itself remains significant at 5% in 14 of 18 models in both conditions.

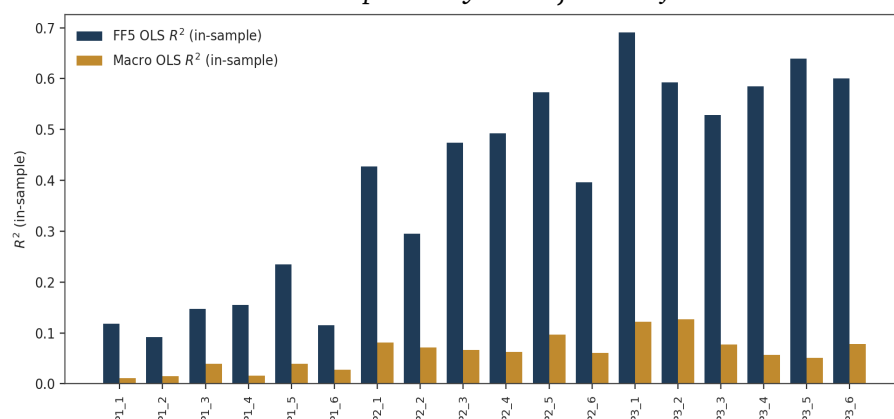
Macroeconomic Pricing Results

The macroeconomic OLS specification is considerably weaker than FF5: average R² across the 18 portfolios is 0.061, versus 0.397 for FF5, and only 9 of 18 models are jointly significant at the 10% level (Table 4; Figure 3). Exchange-rate depreciation is the most consistent predictor, significant at 5% or better in 17 of 18 portfolios with a uniformly negative coefficient, and with larger-magnitude coefficients in the P3 group (–0.008 to –0.012) than the P1 group (–0.002 to –0.005), consistent with greater import-cost and profitability sensitivity among larger, more internationally exposed firms. CPI inflation is positive in 16 of 18 portfolios (significant in 7, concentrated in the P3 group), and industrial-production growth is positive and significant in four portfolios (P2_5, P3_2, P3_5, P3_6), consistent with industrial activity signalling corporate income growth. Oil-price changes and M2 growth are not consistently significant.

Table 4*Full-Sample Macroeconomic OLS Results, 18 Portfolios (Selected)*

Portfolio	ExRate_ret	CPI	IPI_ret	R ²	F p-value
P2_6	−0.0064**	0.0024*	0.0005	0.061	0.112
P3_1	−0.0103***	0.0041**	0.0026*	0.122	0.001
P3_2	−0.0090**	0.0056***	0.0044***	0.127	0.001
P3_6	−0.0115**	0.0046**	0.0039**	0.078	0.035

Note: *** p<0.01, ** p<0.05, * p<0.10. Coefficients for Oil_ret, M2_ret, GDP and PolicyRate omitted for space; full coefficient matrix available from the authors. Full 18-portfolio table available from the authors.

Figure 1*FF5 vs Macroeconomics OLS Explanatory Power, 18 Portfolios*

Machine-Learning Forecasting Results

Table 5 and Figure 4 report out-of-sample performance for Random Forest and ANN using the FF5 predictor set. Random Forest achieves an average out-of-sample R² of 0.645 and mean directional accuracy of 80.3% across the 18 portfolios, with the P3_1 portfolio the strongest individual result (R²=0.905, MDA=94.8%). ANN performs considerably worse on the same predictor set and holdout (average R²=0.055; MDA=68.5%), with only 7 of 18 portfolios showing positive out-of-sample R² and particularly poor generalisation for P1_1 (−0.469), P1_6 (−0.849) and P2_3 (−0.433); this pattern is consistent with the financial machine-learning literature's finding that ensemble tree methods tend to outperform neural networks in medium-sized, monthly-frequency financial datasets (Fallahgoul et al., 2024). A separate rolling-window check (48-month training, 6-month test windows, 2010-2025) for the best-performing P3_1 portfolio finds that Random Forest's out-of-sample predictability is not uniform through time: it is strongest during 2015-2019 and the post-2021 recovery, and weakens markedly around the 2020 COVID-19 shock, indicating that the single 70/30 holdout result, while informative, does not fully characterise the model's stability across sub-periods.

Table 5*Random Forest and ANN Out-of-Sample Performance, FF5 Factors, 70/30 Chronological Split (Selected Portfolios)*

Portfolio	RF R ²	RF RMSE	RF MDA%	ANN R ²	ANN RMSE	ANN MDA%
P1_1	0.335	0.0398	50.0	−0.469	0.0591	51.7
P1_4	0.711	0.0608	87.9	0.314	0.0937	70.7
P2_5	0.806	0.0329	74.1	−0.281	0.0846	60.3
P3_1	0.905	0.0308	94.8	0.544	0.0675	79.3
P3_6	0.855	0.0517	86.2	0.570	0.0890	84.5
Average (18)	0.645	0.0454	80.3	0.055	0.0757	68.5

Note: OOS = out-of-sample. MDA% = mean directional accuracy. Full 18-portfolio table for both models, and Bullish/Bearish sub-sample results, available from the authors.

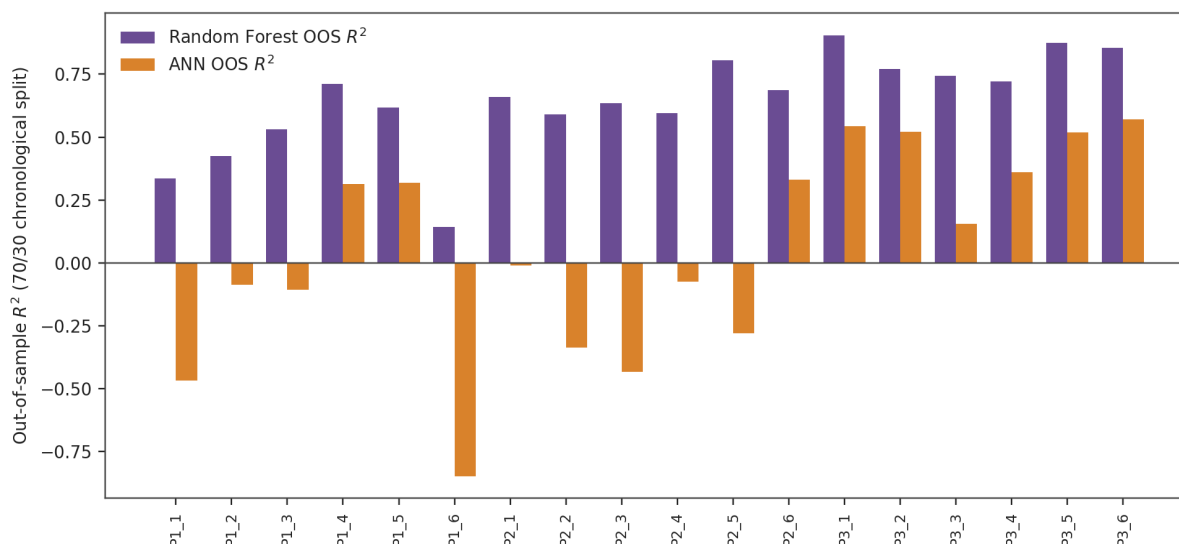
Figure 2*Random Forest vs ANN Out-of-Sample R^2 , FF5 Factors, Full Sample*

Table 6 reports the corresponding results for the macroeconomic predictor set. The models fail to generalize out of sample, with average R^2 being negative for Random Forest (-0.272 overall; -0.976 in Bullish conditions; -0.623 in Bearish conditions) and even lower for ANN (-4.475 overall; -25.329 in Bullish conditions; -8.319 in Bearish conditions), while Random Forest only obtains positive out-of-sample R^2 in one of 18 portfolios (P2_6, $R^2 = 0.018$), and ANN does not obtain any positive out-of-sample R^2 . Directional accuracy provides better insights, with the Random Forest macro model predicting the direction of the return 74.3% of the time in Bullish months and 74.7% of the time in Bearish months, both substantially higher than the 50% random accuracy benchmark, despite the fact that its forecasts are very unreliable.

Table 6*Machine-Learning Performance Summary, Macroeconomic Variables (Average Across 18 Portfolios)*

Model	Market condition	Avg R^2	Avg RMSE	Avg MDA (%)	# Portfolios $R^2 > 0$
Random Forest	Full	-0.272	0.0978	45.6	1 / 18
Random Forest	Bullish	-0.976	0.1012	74.3	2 / 18
Random Forest	Bearish	-0.623	0.0703	74.7	3 / 18
ANN	Full	-4.475	0.1671	53.2	0 / 18
ANN	Bullish	-25.329	0.3118	54.2	0 / 18
ANN	Bearish	-8.319	0.1508	66.1	0 / 18

Note: Negative R^2 indicates out-of-sample prediction worse than the historical mean. This comprehensive negative result is discussed in this paper

Comprehensive Model Comparison

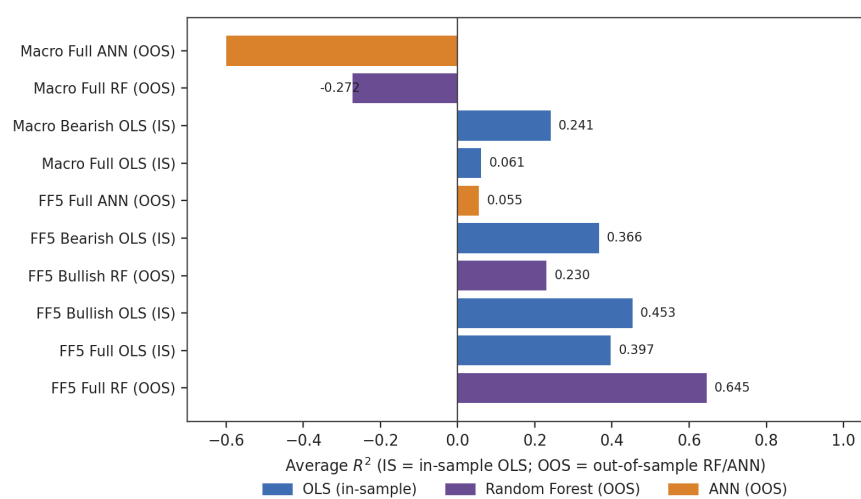
Table 7 and Figure 7 rank all model-factor-set-market-condition combinations by average R^2 , with in-sample OLS and out-of-sample RF/ANN results clearly labelled and never merged into a single ranking. Random Forest with FF5 factors in the full sample achieves the highest out-of-sample R^2 (0.645) among the machine-learning specifications; OLS with FF5 factors provides the strongest interpretable linear benchmark (in-sample $R^2=0.397$, 100% of models jointly significant); and macroeconomic machine-learning models occupy the bottom of the ranking, with negative average R^2 in every condition for both algorithms.

Table 7
Comprehensive Model Comparison, Average Performance Across 18 Portfolios (Ranked by R²)

Rank	Factor set	Market condition	Model	Avg R ²	Avg (%)	MDA	Assessment
1	FF5	Full	Random Forest (OOS)	0.645	80.3		Best OOS model
2	FF5	Full	OLS (IS)	0.397	—		Strong linear baseline
3	FF5	Bullish	OLS (IS)	0.453	—		Highest OLS in-sample fit
4	FF5	Bullish	Random Forest (OOS)	0.230	81.3		Good; limited by small N
5	FF5	Bearish	OLS (IS)	0.366	—		OLS most reliable in Bearish
6	FF5	Full	ANN (OOS)	0.055	68.5		Weak R ² ; retains directional skill
7	Macro	Full	OLS (IS)	0.061	—		Low but statistically significant
8	Macro	Bearish	OLS (IS)	0.241	—		Some condition-specific macro value
9	Macro	Full	Random Forest (OOS)	−0.272	45.6		Overfits macro signal
10	Macro	Full	ANN (OOS)	−4.475	53.2		Poorest overall performer

Note: IS = in-sample (OLS); OOS = out-of-sample, 70/30 chronological split (RF, ANN). OLS and machine-learning R² are reported separately and are not ranked on a single equivalent scale.

Figure 3
Comprehensive Model Comparison, Average R², Across 18 Portfolios

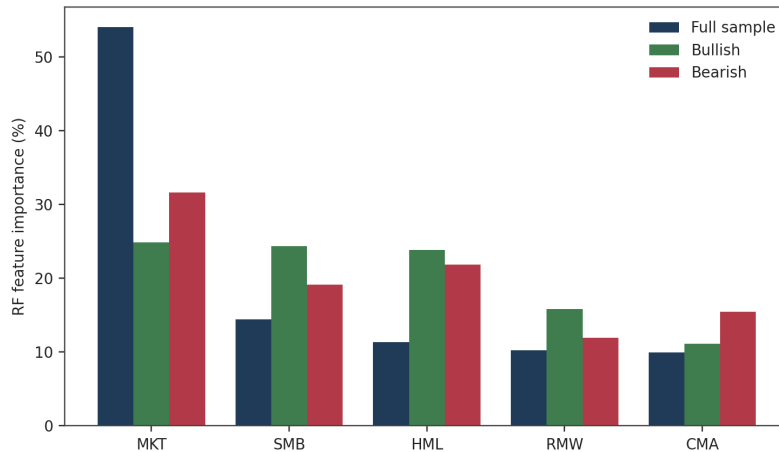


Feature Importance Across Market Conditions

Random Forest feature importance for the FF5 factors (Table 8; Figure 5) shows MKT dominating in the full sample (54.07%) but falling by roughly half in the Bullish condition (24.85%), where SMB (24.33%) and HML (23.82%) become comparably important, suggesting more active size/value style rotation during rising markets; RMW is relatively more important in Bullish conditions (15.86%), consistent with a quality premium during expansions, while CMA is most important in the Bearish condition (15.47%), consistent with a flight-to-quality, investment-conservatism effect during downturns. For macroeconomic variables (Table 9; Figure 6), the exchange rate is the most important predictor across all three conditions (21.80% full; 19.42% Bullish; 19.14% Bearish); oil-price and CPI importance both rise in the Bearish condition (17.88% and 16.35% respectively), consistent with heightened sensitivity to commodity shocks and inflation risk during downturns in a net oil-importing economy; and the policy rate is consistently the least important variable, likely reflecting its high pairwise correlation with CPI (0.797), which reduces each variable's marginal informational contribution when the other is present.

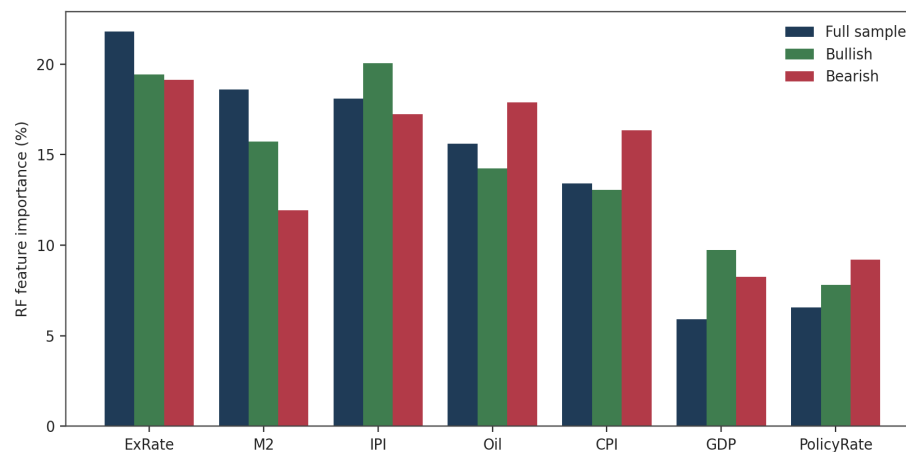
Table 8*RF Feature Importance, FF5 Factors, by Market Condition (Average Across 18 Portfolios, %)*

Factor	Full sample	Bullish	Bearish
MKT	54.07	24.85	31.61
SMB	14.40	24.33	19.14
HML	11.32	23.82	21.83
RMW	10.28	15.86	11.95
CMA	9.93	11.15	15.47

Figure 4*RF Feature Importance, FF5 Factors, by Market Condition***Table 9***RF Feature Importance, Macroeconomic Variables, by Market Condition (Average Across 18 Portfolios, %)*

Factor	Full sample	Bullish	Bearish
ExRate_ret	21.80	19.42	19.14
M2_ret	18.61	15.72	11.92
IPI_ret	18.08	20.05	17.24
Oil_ret	15.61	14.23	17.88
CPI	13.42	13.04	16.35
GDP	5.92	9.73	8.26
PolicyRate	6.56	7.80	9.21

Note: Values sum to 100% within each column. Importance is a measure of predictive association within the fitted Random Forest, not a causal contribution to returns.

Figure 5*RF Features Importance, Macroeconomic Variables, by Market Condition*

As far as economic value is concerned, a directional accuracy of 80.3% on average implies that in about four of every five months where the Random Forest FF5 was evaluated out-of-sample, it correctly signed the return direction of the portfolios in question. In turn, RMSE of 0.0454 equates to 4.5 percentage points of monthly return forecasting error. This information might be useful for studying market timing and risk management; nevertheless, the analysis itself does not incorporate any consideration of transaction costs, portfolio turnover, adjusted returns, or trading strategy implementation.

Discussion

FF5 Factor Pricing in the PSX

The OLS findings offer compelling proof of the empirical significance of the FF5 model in the PSX, as the MKT is statistically significant in all portfolios, and the monotonic rise in MKT from P1 to P3 is in line with the theoretical expectation that market-sensitive portfolios will have higher risk exposure. The uniformly negative SMB loadings, however, indicate an inverted or absent size premium relative to much of the developed-market literature, consistent with a portfolio-construction methodology weighted toward larger, more liquid PSX constituents (Sha & Gao, 2019), and the predominantly negative HML loadings point to a growth rather than value tilt over this sample period.

Macroeconomic Factor Pricing and the APT

The macroeconomic results support the APT's qualitative prediction that macroeconomic state variables are priced, but with far lower explanatory power than firm-level factors (average $R^2=0.061$ versus 0.397). The consistent, negative exchange-rate effect is consistent with the theoretical channel through which currency depreciation raises import costs, compresses real corporate profitability, and increases the equity risk premium in an economy reliant on imported inputs and energy; the positive CPI relationship may reflect nominal-return adjustment during inflationary periods or the coincidence of inflationary episodes with fiscally accommodative conditions that temporarily supported equity valuations. The overall pattern is consistent with prior Pakistani evidence and the broader international literature reporting weak macroeconomic return predictability in the cross-section.

Machine Learning, Nonlinearity, and the Negative Macro-ML Result

Random Forest's out-of-sample advantage over OLS-style linear fit and over ANN in the FF5 setting is consistent with the interpretation that PSX returns contain nonlinear interactions and threshold effects among MKT, SMB, HML, RMW and CMA that an ensemble tree method can exploit but a single linear coefficient cannot; this reading is consistent with prior evidence that Random Forest outperforms linear and neural-network benchmarks in medium-sized financial datasets (Breiman, 2001; Fallahgoul et al., 2024; Yao et al., 2022). The comprehensive failure of the macroeconomic machine-learning models to generalise out of sample is, however, an equally important and theoretically informative finding, not merely a null result to be set aside: theoretical relevance, established through the OLS results, does not imply out-of-sample forecasting stability. Plausible explanations — offered as interpretations rather than proven mechanisms — include structural breaks associated with successive monetary-policy, exchange-rate and fiscal regimes over 2010-2025, a low signal-to-noise ratio in monthly macro-return relationships, and the small number of effective degrees of freedom available to the machine-learning models once the sample is split chronologically. This pattern echoes the broader empirical finance literature on the limits of macroeconomic return predictability (Goyal & Welch, 2008) and suggests that, in a volatile emerging market, macroeconomic information may be better incorporated through factor-model regressions than through direct machine-learning forecasts of returns.

State Dependence and the Adaptive Market Hypothesis

The variation in factor loadings, model fit, feature importance and machine-learning predictability across Bullish, Bearish and Neutral conditions is consistent with the state-dependent mechanism emphasised by the

Adaptive Market Hypothesis (Lo, 2004): MKT's dominant full-sample RF importance (54.07%) recedes sharply within both classified conditions in favour of SMB, HML, RMW and CMA, and macro-feature importance reweights toward oil and CPI in Bearish periods. We do not read these results as proving the AMH; rather, the variation in loadings, fit, feature importance and predictive performance across market conditions is consistent with the state-dependent pricing and evolving-efficiency mechanisms the AMH describes, within the specific PSX sample and classification rule used here.

Statistical versus Economic Significance

Three distinct claims must not be conflated. A significant OLS coefficient establishes a statistical association within the specified linear model; a positive out-of-sample R^2 or high MDA establishes forecasting skill relative to a historical-mean or random benchmark on the specific holdout tested; and neither establishes that a trading strategy built on these forecasts would be profitable once transaction costs, market impact, portfolio turnover and implementation constraints are taken into account, none of which are estimated in this study. The 80.3% Random Forest MDA is a meaningful statistical result; it is not evidence of achievable investment profitability, and we explicitly avoid that inference.

Practical, Managerial and Policy Implications

For portfolio managers and analysts, the Random Forest FF5 results and the market-condition-specific feature-importance rankings may serve as an exploratory research tool and a benchmark for further model development and monitoring, rather than as a validated basis for live investment decisions; validation through additional forecast windows, transaction-cost analysis and out-of-sample testing beyond the single holdout used here would be required before any such recommendation. For financial-market regulators and policymakers, the consistent negative exchange-rate relationship and the heightened importance of oil and inflation variables during bearish conditions indicate areas warranting continued monitoring for financial-stability purposes, but the associations are observational, not causal, and should not be read as prescriptions for specific monetary, exchange-rate or fiscal actions.

Theoretical Contribution

The theoretical contribution is the integration, within a single empirical architecture, of APT-motivated macroeconomic risk, FF5 firm-level systematic risk, and AMH-motivated state dependence, with machine learning serving as the flexible estimation method used to test whether these theoretically grounded relationships are nonlinear, interactive and market-condition dependent. The study does not propose a new asset-pricing theory; its contribution is evidence that predictability in the PSX is jointly predictor-set dependent, model dependent and market-condition dependent, within the sample, holdout design and specifications examined.

Conclusions

This study examined whether asset-pricing relationships and return predictability in the Pakistan Stock Exchange vary across market conditions, and whether machine-learning methods capture this state dependence more flexibly than linear models. Using 18 PSX portfolio-return series, two theoretically motivated predictor sets (FF5 factors; APT-motivated macroeconomic variables), three modelling approaches (OLS, ANN, Random Forest) and an explicit Bullish/Bearish/Neutral classification, several findings emerge. The FF5 model has meaningful explanatory relevance in the PSX, with MKT significant across all 18 portfolios and average in-sample R^2 of 0.397; macroeconomic predictors are statistically related to returns — particularly the exchange rate, CPI and industrial production — but explain much less of their variation (average $R^2=0.061$). Under a 70/30 chronological holdout, Random Forest extracts substantial predictive structure from the FF5 factors (average out-of-sample $R^2=0.645$; MDA=80.3%), while the ANN is markedly more sensitive to the limited monthly sample ($R^2=0.055$). Macroeconomic machine-learning models do not generalise out of sample, with negative average R^2 for both algorithms across all market conditions. Factor

loadings and feature importance shift materially between Bullish, Bearish and Neutral conditions, a pattern consistent with state-dependent asset pricing and the Adaptive Market Hypothesis.

These results are conditional on the sample period, the pre-specified model architectures, the single chronological train/test split, and the classification rule used to define market conditions; they should not be read as establishing universal machine-learning superiority over linear models — OLS remains the more reliable and interpretable tool for inference, particularly in the small Bullish and Bearish sub-samples — nor as evidence that the Random Forest forecasts are, or would be, profitable in live trading, since transaction costs, turnover and risk-adjusted implementation were not evaluated. The market-state classification is contemporaneous rather than strictly ex-ante, and a fair, same-test-period OLS forecasting benchmark was not available for direct comparison against ANN and Random Forest; both points qualify the strength of causal or forecasting claims that can be drawn and are identified as priorities for further work. Within these bounds, the evidence supports a central, more nuanced conclusion than any single-model ranking: PSX return predictability depends jointly on which predictors are used, which model estimates the relationship, and which market condition prevails, and machine learning is best understood as a flexible complement for exploring this state dependence rather than a replacement for theory-grounded factor models or a validated investment tool.

References

- Armano, G., Marchesi, M., & Murru, A. (2005). A hybrid genetic-neural architecture for stock indexes forecasting. *Information Sciences*, 170(1), 3–33. <https://doi.org/10.1016/j.ins.2003.03.023>
- Bagnara, M., Goodell, J. W., Ortmann, R., & Staudinger, M. (2024). Asset pricing and machine learning: A critical review. *Journal of Economic Surveys*, 38(1), 27–56. <https://doi.org/10.1111/joes.12532>
- Bradrania, R., & Pirayesh Neghab, D. (2022). State-dependent asset allocation using neural networks. *The European Journal of Finance*, 28(11), 1125–1150. <https://doi.org/10.1080/1351847X.2021.1960404>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Cao, Q., Parry, M. E., & Leggio, K. B. (2011). The three-factor model and artificial neural networks: Predicting stock price movement in China. *Annals of Operations Research*, 185(1), 25–44. <https://doi.org/10.1007/s10479-009-0618-0>
- Carvalho, A., & Ribeiro, T. (2008). Do artificial neural networks provide better forecasts? Evidence from Latin American stock indexes. *Latin American Business Review*, 9(3), 92–110. <https://doi.org/10.1080/10978520802035463>
- Chiah, M., Chai, D., Zhong, A., & Li, S. (2016). A better model? An empirical investigation of the Fama–French five-factor model in Australia. *International Review of Finance*, 16(4), 595–638. <https://doi.org/10.1111/irfi.12099>
- Chen, L., Pelger, M., & Zhu, J. (2024). Deep learning in asset pricing. *Management Science*, 70(2), 714–750. <https://doi.org/10.1287/mnsc.2023.4695>
- Chen, S.-J., & Jordan, B. D. (1993). Some empirical tests in the arbitrage pricing theory: Macro variables vs. derived factors. *Journal of Banking & Finance*, 17(1), 65–89. [https://doi.org/10.1016/0378-4266\(93\)90080-W](https://doi.org/10.1016/0378-4266(93)90080-W)
- Cheng, T., Jiang, S., Zhao, A. B., & Zhao, J. (2025). Is machine learning a necessity? A regression-based approach for stock return prediction. *Journal of Empirical Finance*, 81, Article 101598. <https://doi.org/10.1016/j.jempfin.2025.101598>
- Connor, G., & Korajczyk, R. A. (1995). The arbitrage pricing theory and multifactor models of asset returns. In R. Jarrow, V. Maksimovic, & W. T. Ziemba (Eds.), *Handbooks in operations research and management science: Finance* (Vol. 9, pp. 87–144). Elsevier. [https://doi.org/10.1016/S0927-0507\(05\)80048-9](https://doi.org/10.1016/S0927-0507(05)80048-9)
- Fallahgoul, H., Franstianto, V., & Lin, X. (2024). Asset pricing with neural networks: Significance tests. *Journal of Econometrics*, 238(1), Article 105574. <https://doi.org/10.1016/j.jeconom.2023.105574>
- Fama, E. F., & French, K. R. (2017). International tests of a five-factor asset pricing model. *Journal of Financial Economics*, 123(3), 441–463. <https://doi.org/10.1016/j.jfineco.2016.11.004>
- Fang, M., & Taylor, S. (2021). A machine learning based asset pricing factor model comparison on anomaly portfolios. *Economics Letters*, 204, Article 109919. <https://doi.org/10.1016/j.econlet.2021.109919>
- Fatima, S., & Hussain, G. (2008). Statistical models of KSE100 index using hybrid financial systems. *Neurocomputing*, 71(13–15), 2742–2746. <https://doi.org/10.1016/j.neucom.2007.11.044>
- Goyal, A., & Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. *The Review of Financial Studies*, 21(4), 1455–1508. <https://doi.org/10.1093/rfs/hhm014>
- Gu, S., Kelly, B., & Xiu, D. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 33(5), 2223–2273. <https://doi.org/10.1093/rfs/hhaa009>
- Hanauer, M. X., & Kalsbach, T. (2023). Machine learning and the cross-section of emerging market stock returns. *Emerging Markets Review*, 55, Article 101022. <https://doi.org/10.1016/j.ememar.2023.101022>
- Haqqani, K., & Aleem, M. (2020). Testing liquidity augmented Fama-French five-factor model in Pakistan Stock Exchange. *Global Economics Review*, 5(1), 255–265. [https://doi.org/10.31703/ger.2020\(V-I\).21](https://doi.org/10.31703/ger.2020(V-I).21)

- He, Z., Zhu, J., & Zhu, X. (2015a). Dynamic factors and asset pricing: International and further U.S. evidence. *Pacific-Basin Finance Journal*, 32, 21–39. <https://doi.org/10.1016/j.pacfin.2015.02.002>
- He, Z., Zhu, J., & Zhu, X. (2015b). Multi-factor volatility and stock returns. *Journal of Banking & Finance*, 61(S2), S132–S149. <https://doi.org/10.1016/j.jbankfin.2015.09.013>
- Idowu, P. A., Osakwe, C., Kayode, A. A., & Adagunodo, E. R. (2012). Prediction of stock market in Nigeria using artificial neural network. *International Journal of Intelligent Systems and Applications*, 4(11), 68–74. <https://doi.org/10.5815/ijisa.2012.11.08>
- Jan, M. N. (2019). *Asset pricing and artificial neural networks: A case of Pakistan's equity market* [Doctoral dissertation, Capital University of Science and Technology].
- Kazak, H., Kumar, S., Gündüz, M. A., & Akcan, A. T. (2025). Metaheuristic-optimized ANFIS and ANN models for stock price forecasting: Evidence from the Borsa Istanbul 100 index. *Discover Artificial Intelligence*, 5, Article 272. <https://doi.org/10.1007/s44163-025-00395-6>
- Kelly, B. T., Malamud, S., & Zhou, K. (2024). The virtue of complexity in return prediction. *The Journal of Finance*, 79(1), 459–503. <https://doi.org/10.1111/jofi.13298>
- Khan, N., Zada, H., Ahmed, S., Shah, F. A., & Jan, S. (2023). Human capital-based four-factor asset pricing model: An empirical study from Pakistan. *Heliyon*, 9(5), Article e16328. <https://doi.org/10.1016/j.heliyon.2023.e16328>
- Liang, D., Lu, C.-C., Tsai, C.-F., & Shih, G.-A. (2024). Forecasting corporate financial performance with deep learning and interpretable ALE method: Evidence from China. *Journal of Forecasting*, 43(7), 2540–2571. <https://doi.org/10.1002/for.3138>
- Lo, A. W. (2004). The adaptive markets hypothesis: Market efficiency from an evolutionary perspective. *The Journal of Portfolio Management*, 30(5), 15–29. <https://doi.org/10.3905/jpm.2004.442611>
- Munir, T., Saeed, M., & Aziz, B. (2023). Forecasting the Pakistan's leading stock exchange during COVID-19 using machine learning. *Research Square*. <https://doi.org/10.21203/rs.3.rs-2498599/v1>
- Nyberg, H. (2013). Predicting bear and bull stock markets with dynamic binary probit models. *Journal of Banking & Finance*, 37(8), 2952–2963. <https://doi.org/10.1016/j.jbankfin.2013.05.008>
- Priestley, R. (1996). The arbitrage pricing theory, macroeconomic and financial factors, and expectations generating processes. *Journal of Banking & Finance*, 20(5), 869–890. [https://doi.org/10.1016/0378-4266\(95\)00035-6](https://doi.org/10.1016/0378-4266(95)00035-6)
- Raza, H., & Akhtar, Z. (2024). Predicting stock prices in the Pakistan market using machine learning and technical indicators. *Modern Finance*, 2(2), 46–63. <https://doi.org/10.61351/mf.v2i2.167>
- Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, 13(3), 341–360. [https://doi.org/10.1016/0022-0531\(76\)90046-6](https://doi.org/10.1016/0022-0531(76)90046-6)
- Sha, Y., & Gao, R. (2019). Which is the best: A comparison of asset pricing factor models in Chinese mutual fund industry. *Economic Modelling*, 83, 8–16. <https://doi.org/10.1016/j.econmod.2019.09.016>
- Ullah, R., Jan, M. N., & Tahir, M. (2025). Unveiling the optimal factor model in Pakistan: A machine learning approach using support vector regression and extreme gradient boosting algorithms. *Future Business Journal*, 11, Article 137. <https://doi.org/10.1186/s43093-025-00560-4>
- Yao, H., Xia, S., & Liu, H. (2022). Six-factor asset pricing and portfolio investment via deep learning: Evidence from Chinese stock market. *Pacific-Basin Finance Journal*, 76, Article 101886. <https://doi.org/10.1016/j.pacfin.2022.101886>
- Zaremba, A., & Szczygielski, J. J. (2019). Limits to arbitrage, investor sentiment, and factor returns in international government bond markets. *Economic Research-Ekonomska Istraživanja*, 32(1), 1727–1743. <https://doi.org/10.1080/1331677X.2019.1638286>
- Zegadło, P. (2022). Identifying bull and bear market regimes with a robust rule-based method. *Research in International Business and Finance*, 60, Article 101603. <https://doi.org/10.1016/j.ribaf.2021.101603>