ADVANCE SOCIAL SCIENCE ARCHIVE JOURNAL
Available Online: https://assajournal.com
Vol. 05 No. 01. Jan-March 2026.Page# 7495-7520
Print ISSN: 3006-2497 Online ISSN: 3006-2500
Platform & Workflow by: Open Journal Systems
https://doi.org/10.5281/zenodo.22867381

Financial Stability and Risk Resilience in Pakistan's Islamic and Conventional Banks: Evidence from Dynamic Econometrics and AI Forecasting

Ghulam Nabi (Corresponding Author)
College of Statistical Sciences, University of the Punjab, Lahore, Pakistan.
gnabi.stat@pu.edu.pk
Shamsa Yousaf
College of Statistical Sciences, University of the Punjab, Lahore, Pakistan.
shamsauosaf@gmail.com
Aqsa Shafi
College of Statistical Sciences, University of the Punjab, Lahore, Pakistan.
aqsas833@gmail.com

ABSTRACT
This article looks at the impact of credit, liquidity and solvency risk on profitability based financial stability in Pakistan's dual banking system and compares the findings of dynamic econometric models with AI forecasting models for complementary evidence for risk surveillance. Using a balanced annual panel of 21 Pakistani banks from 2010 to 2024, including 18 conventional and three fully fledged Islamic banks, the study uses return on assets (ROA) as the stability signal. The non-performing loans (NPL) and the ratio of liquid assets to deposits (LR) and that of liabilities to assets (SER) measure the credit, liquidity and solvency risks respectively, while GDP growth and inflation capture the macroeconomic environment. Analysis includes descriptive statistics, correlations, lag-order selection, Johansen cointegration, Panel VECM, Difference GMM and AI forecasting models (Random Forest, MLP, Hybrid RF-MLP, LSTM and GRU). The analysis reveals that the average ROA, NPLs and ROA volatility of Islamic banks are better than those of conventional banks, in addition to having a better solvency position. The most constant risk channel is credit risk. The findings of VECM support that the profitability indicators are in long run equilibrium, and the findings of GMM show the persistence and the solvency related effects with Islamic banks. The results indicate that GRU works best with conventional banks, whereas LSTM and Hybrid RF-MLP are more effective with Islamic banks using error measures. The results are relevant for risk surveillance of banks.

***Keywords**: Islamic banking; Conventional banking; Financial stability; Credit risk; Liquidity risk; Solvency risk; Panel VECM; AI forecasting.*

1. Introduction

Banks' stability is a crucial enabler for sustainable economic development, as they catalyse savings, distribute credit resources, and facilitate payment mechanisms and moderate financial shocks. In modern banking research, stability is no longer regarded as a straightforward result of profitability. It is now recognized as the outcome of the interplay of risk channels, mainly credit quality and liquidity capacity and solvency strength. Bank fragility typically occurs as the result of a combination of risks. Rather, deterioration in asset quality is more likely to cause decreases in profitability, dampen funding confidence, create liquidity stress, and strain capital buffers. It resurfaced during the global financial crisis when credit losses hurt bank balance sheets and drove a disruption in liquidity markets (Brunnermeier, 2009) and again during COVID-19 when macroeconomic shocks impacted bank profitability, asset quality and risk-taking (Demirgüç-Kunt et al., 2021; Ghenimi, Chaibi & Omri, 2024). These interactions are especially relevant in emerging markets as volatility in inflation, exchange-rate pressures, sovereign-risk exposure, and shallowness in financial markets and regulatory transitions can quickly drive-up bank-level risks to a higher level of system fragility.

The time is opportune to look at the issue of banking risk resilience in Pakistan. The country has a dual banking system comprising conventional and Islamic banking sector which operates under the dual regulatory framework of State Bank of Pakistan. Conventional banks work on the basis of interest intermediation, while Islamic banks are based on Shariah principles, which do not allow the practices of riba, asset-based finance, trade, and risk sharing. The differences found in these institutions indicate that the flow of liquidity, credit and solvency shocks may not impact both Islamic and conventional banks in the same way. Islamic banks can benefit from asset-based financing, contractual discipline and tighter financial and real economic activity relationship. However, they may also encounter liquidity-management limitation due to the fact that the liquidity management instruments, the hedging facilities and the interbank facilities, which are available to Shariah-compliant banks, are less developed than those available to conventional banks. Conventional banks, on the other hand, generally enjoy greater access to liquidity instruments based on interest, as well as a well-developed money market, but they may be more vulnerable to credit-cycle risk as intermediated, through lending, and competitive. Relying on a lump-sum classification of all Pakistani banks may mask significant variations in business model, risk transmission, and rate of adjustment and predictive stability.

With the advent of Islamic banking as an important part of the national financial system, the Pakistan context is now gaining significance. According to the latest statistics from the State Bank of Pakistan (SBP), Islamic banking assets stood at PKR 12,345 billion by June 2025, and deposits stood at PKR 9,533 billion. Islamic banking assets amounted to 20.7 per cent of the banking industry, while Islamic banking deposits were 25.5 per cent of the total banking industry deposits (SBP, 2025b). Meanwhile, Pakistan's banking sector is facing macroeconomic and structural stresses, such as inflation volatility, monetary policy changes, financing requirements of the public sector, asset quality issues and more. SBP's latest financial stability reviews keep highlighting that the banking sector is resilient, and continues to stress test and supervise the sector in adverse macro-financial conditions (SBP, 2025a). In the face of increasing Islamic banking share, the resilience of the overall banking system and the continued macro-financial stress in the country, Pakistan is a suitable country to examine the risk resilience of Islamic and conventional banks.

There are valuable but partial guidelines in the literature. While much attention is given to credit risk and liquidity risk, solvency risk is comparatively less integrated. Solvency ultimately determines whether a bank will be able to withstand adverse events and remain in operation. This has been tackled by recent studies. Akram and Hushmat (2024) demonstrate that there is significant relationship between liquidity creation and solvency risk in Islamic banks and conventional banks of Pakistan and Malaysia. Using a Panel Vector Error Correction Model, Peykani et al. (2025) show that there is a dynamic link between the three types of risk in emerging-economy banks: liquidity risk, credit risk and solvency risk. But the studies either failed to focus particularly on Pakistan's dual banking system or lacked the explanatory econometrics and the use of AI based forecasting. Similarly, Islamic-conventional comparative studies focus primarily on profitability and efficiency, and stability/creditor risk differences (

Beck, Demirgüç-Kunt & Merrouche, 2013; Kabir, Worthington & Gupta, 2015; Hassan, Khan & Paltrinieri, 2019; Ghenimi, Chaibi & Omri, 2024), but they do not jointly model liquidity risk, credit risk and solvency risk within a single long-run and short-run framework.

Liquidity creation is another recent stream of research that investigates the existence of the differences between Islamic and conventional banks, even under similar regulatory frameworks, with a focus on liquidity risk and bank-type differences (Akram & Hushmat, 2026). This underlines the importance of assessing Pakistan banking system on a dual banking system basis, which requires a risk integrated and model integrated approach.

The second and smaller drawback has to do with methodology. Econometric models are valuable for explaining the relationships, testing the long run equilibrium and estimating the adjustment behaviour. But they are not used chiefly for predictions. Nonlinear pattern recognition and forecasting with machine-learning and deep-learning models are more effective, but tend to lack economic interpretation. In recent years, credit and bank-risk data classification research with AI has been conducted, with an increasing number of studies employing Random Forests, neural networks, gradient-boosting models and hybrid models. (Sugozu, Verberi & Yasar, 2025; Gafsi, 2025). AI-based bank-risk studies are, however, generally not combined with econometric stability analysis. This separation is very important due to the need of the regulators and bank managers for explanatory evidence and predictive warning signals. They must understand and be able to identify what risks are important, how banks react following an event, and if there is still some information that is available at the bank level that can predict future bank outcomes with reasonable precision.

This article addresses these limitations by building a single explanation and prediction model of Pakistan's dual banking system. The explanatory layer considers Panel VECM to explore the adjustment process in the short-run and long-run equilibrium for the financial stability, liquidity risk, credit risk and solvency risk. Then, dynamism and potential endogeneity of bank profitability are dealt with by using Difference GMM. The predictive layer compares the performance of the following models for ROA prediction: Random Forest, MLP, Hybrid RF-MLP, LSTM and GRU models. The study is theoretically relevant by connecting financial instability theory, Basel oriented prudential logic and dual banking theory and methodologically it combines dynamic econometrics with AI forecasting which is relevant for practice, as it provides evidence for risk sensitive supervision, capital planning, liquidity management and early-warning system design.

The article discusses four research questions. To begin with, what are the liquidity, credit and solvency risks for bank stability in Pakistan? Second, are there differences in the risk resilience and adjustment behaviour of Islamic and conventional banks? Third, do the evidence from the econometric and AI-based models complement each other in the context of banking-risk surveillance? Fourthly, what are the better predictive models for forecasting ROA under the dual banking system in Pakistan?

In this article, the authors add four points to the banking-risk literature. First, it offers an integrated perspective on financial stability by modelling credit, liquidity and solvency risks simultaneously, rather than as banking ratios. Secondly, it compares Islamic and conventional banks by presenting the analysis of both banks in the same data period and methodology. Third, there is a predictive element that combines dynamic econometric evidence and the predictive performance of machine-learning and deep-learning models. Fourth, it provides practical benefit by providing bank-specific implications for supervision and risk monitoring and managerial decision making.

2. Theoretical Background and Hypothesis Development

2.1 Financial instability and integrated risk channels

This study has a theoretical foundation starting from the financial instability theory. During a seeming period of stability, financial systems can grow more fragile, as banks and borrowers become more leveraged, and confidence in their ability to fulfill obligations is reduced. In banking, this applies to the relationship between credit quality, liquidity and financial strength or solvency. Once assets deteriorate, profits drop, liquidity gets tight and capital buffers could be used up. It is a transition from apparent stability to financial fragility, in line with the evidence from the post-crisis period that shocks are cascaded through the balance sheets of banks, funding markets and the expectations of macro-financial agents (Brunnermeier, 2009; Demirgüç-Kunt et al., 2021).

A second lens is prudential regulation based on Basel. Basel III was developed to toughen up the bank regulation, supervision and risk management in the wake of a global financial crisis. It's based on the logic that, besides credit-risk management, there needs to be adequate liquidity, capital strength, and leverage discipline. Capital losses due to credit losses may make a bank with sufficient liquidity but poor asset quality look unstable. Likewise, a bank with satisfactory capital ratios can put the strain on the institution if the funding needs for the short term become scarce. It is therefore important to consider liquidity, credit and solvency risks together (Basel Committee on Banking Supervision, 2011, 2017).

Solvency is particularly significant because it is the last line of defence against a bank's failure. Liquidities problems can be short-term but solvencies problems are structural. The evidence now shows that these are a joint-risk. Combined liquidity-solvency stress testing is stressed by Cont, Kotlicki and Valderrama (2020). In banking systems, De Bandt, Lecarpentier and Pouvelle (2021) demonstrate that there is an interaction between liquidity and solvency requirements. Peykani et al. (2025) show that there are dynamic relationships between liquidity and credit and solvency risks in emerging economy banks. In light of the above, the present article focuses on the joint behaviour of liquidity risk, credit risk and solvency risk in Pakistan, which is supported by these studies.

2.2 Dual banking and risk-resilience differences

The explanation of the differences in risky behaviour between Islamic and conventional banks can be attributed to dual banking theory. Islamic banks operate under the principles of Shariah which do not allow taking interest and ensure that financing is tied to real assets, trade transactions or risk-sharing arrangements. Theoretically, this could help lessen speculative exposure, enhance asset discipline, and bring financial claims closer to real economic activity. But Islamic banks can also be subjected to liquidity management restrictions as the development of sharia-compliant money-market instruments, hedging instruments and lender oflast-resort facilities are not as advanced as conventional money-market instruments, hedging instruments and lender of last-resort facilities, respectively. In the face of generally stable financial soundness indicators Islamic Financial Services Board has reiterated the significance of enhancing liquidity and risk management skills in Islamic finance (IFSB, 2025).

Conventional banks are based on an interest-based intermediation system, where interest rates are used to price deposits and loans. The structure provides conventional banks with more options for money-market instruments and interbank liquidity facilities, but could also increase their sensitivity to interest rates, their maturity mismatch and their credit-cycle amplification. Thus, the conventional model could lead to quicker shock adaptation, but this quicker shock adaptation does not necessarily mean greater resilience. These differences are theoretical and call for distinct estimations of Islamic and conventional banks instead of a pooled estimation.

This comparative argument has been confirmed by empirical research. Beck, Demirgüç-Kunt and Merrouche (2013) demonstrate that Islamic and conventional banks have different business models, efficiency and stability. Kabir, Worthington and Gupta (2015) argue that the exposure to credit risk may vary between banking systems. Hassan, Khan and Paltrinieri (2019) find distinct differences between liquidity risk, credit risk and stability in Islamic and conventional banks. The recent studies conducted by Ghenimi, Chaibi and Omri (2024) indicate that the response of Islamic banks and conventional banks to the COVID-19 crisis is different in the MENA region. Akram and Hushmat (2024) also demonstrate the significant relationship between liquidity creation and solvency risk in both Islamic and conventional banks in Pakistan and Malaysia.

2.3 Liquidity risk, credit risk and solvency risk

Liquidity risk is the risk of the bank incurring unacceptable losses from having to meet some of their short-term obligations. An increase in liquid-asset holdings could enhance resilience as it will help banks to be able to meet the demands for withdrawals, honour contracts and withstand some funding shocks. The liquidity-profitability relationship, however, may not be positive and unambiguous. When there is too much liquidity, ROA will suffer, as there are likely to be lower returns in liquid assets than in loans and long-term investments. Thus, under conditions of stress, liquidity can help to stabilize the banks, but in the normal period, it can lower profits. The impact could vary by the type of bank, as Islamic banks are more likely to have a more limited choice of sharia-compliant liquidity products, whereas conventional banks are more likely to have access to conventional money markets.

Credit risk is one of the most direct means of which banks lose profitability and stability. Non-performing loans decrease interest or financing income, raise provisioning expenses and degrade the quality of the balance sheet. There can also be impacts on liquidity due to credit deterioration as well, as delayed repayments means reduced cash inflows. Inflation, exchange-rate pressure, macro-economic instability and borrower's vulnerability are often factors in emerging economies that exacerbate credit risk. Credit risk has been repeatedly pointed out as a major factor influencing the performance and stability of banks by prior studies (Ghenimi et al., 2017; Kabir, Worthington & Gupta, 2015; Abdelaziz et al., 2022; Naili & Lahrichi, 2022; Ben-Ahmed et al., 2023; Ghenimi, Chaibi & Omri, 2024).

Solvency risk is the risk that a bank might not be able to pay its long-term obligations due to an excessive amount of liabilities or a lack of capital to absorb losses. An increase in NPLs lowers profitability and capital buffers, and liquidity stress might lead to asset sales and a decline in balance-sheet strength. Oino (2021) associates' solvency with credit risk, liquidity risk, regulatory capital and economic stability. Sakouvogui (2020) explores the key factors of bank liquidity and solvency risks, and emphasizes how these are relevant for bank resilience. According to Peykani et al. (2025) the dynamic relationship between solvency risk and liquidity and credit risk is demonstrated by Panel VECM.

2.4 Econometric explanation and AI-based predictive surveillance

This study's theoretical argument is more of a dynamic rather than a static one. Banking risks change over time. A credit quality shock can impact the profitability of the current period, and can also have an impact on future liquidity and solvency conditions. Likewise, the impacts of a liquidity shock may be indirect on the capacity to lend, on income generation and on capital adequacy. It is therefore appropriate to use the Panel VECM in this case, as it is possible to separate the concepts of long-term equilibrium and short-term adjustment. Another consideration is relevance of Difference GMM: profitability of banks is persistent and can be influenced by endogeneity and omitted heterogeneity and lagged performance.

The methodological layer added by AI-based forecasting is complementary. Econometric models: trying to explain relations; Machine-learning models: trying to measure the predictive usefulness. Random Forest is good for tabular data, and it averages multiple decision trees, which reduces variance from the overfitting of individual trees due to bootstrap sampling and feature randomness (Breiman, 2001). Nonlinearity can exist between financial ratios, which can be captured by models based on neural network. The LSTM model is tailored to incorporate temporal dependencies via memory cells and gates (Hochreiter & Schmidhuber, 1997), whereas GRU's have a simpler recurrent structure and can yield better performance in cases of smaller samples (Cho et al., 2014). But in annual banking panels with only the points of time mentioned, a deep-learning model does not always beat simpler and more interpretable models. Predictive failure can thus be informative, indicating that the information provided in the bank can be noisy, heterogenous or sparse for highly parameterised models.

This is a cautious but positive approach to AI in banking, backed by recent studies. Petropoulos et al. (2020) demonstrated that machine-learning methods can be used to enhance the prediction of bank insolvency. Tavana et al. (2018) show that neural networks and Bayesian networks can detect factors that lead to a liquidity risk. Machine-learning approaches are used by Gafsi (2025) to predict credit risk for participation and conventional banks. In Turkish participation banks and conventional banks, Sugozu, Verberi and Yasar (2025) investigate the credit risk of the banks by applying machine learning algorithms and explainable-AI techniques. However, there is limited research that integrates dynamic econometric risk modelling and AI forecasting in the dual banking system in Pakistan.

2.5 Hypothesis development

H1: Credit risk has a negative and significant effect on bank financial stability, and the effect differs between Islamic and conventional banks.

H2: Liquidity position has a significant effect on bank financial stability, but its short-run influence differs between Islamic and conventional banks.

H3: Solvency risk significantly affects long-run equilibrium and profitability adjustment in both Islamic and conventional banks.

H4: Islamic and conventional banks differ significantly in risk resilience, reflected in profitability, credit quality, solvency strength and adjustment behaviour.

H5: Econometric and AI-based models provide complementary evidence for banking-risk surveillance, but simpler or hybrid machine-learning models may outperform deep-learning models in small annual banking panels.

3. Data and Variable Measurement

The study employs annual data of 21 banks of Pakistan for a 15-year period from 2010 to 2024. The final balanced panel contains 315 observations covering 18 conventional banks for bank-year observations and three fully fledged Islamic banks for 45 observations. The period covers normal years, the COVID-19 shock and the recent macroeconomic stress period in Pakistan. The variables at the bank level are taken from audited financial statements and macroeconomic variables from public sources.

The sample selection follows three criteria: full-period data availability, continuity of reporting and completeness of the required variables. The data audit shows that there are no missing observations across the variables for the main empirical analysis and that each bank makes 15 observations each year of the other variables. The smaller size of the Islamic banking sample is thus a structural property of the well-developed population of Islamic banks in Pakistan and not a data cleaning problem. This restriction is mitigated by the separate estimation, chronological validation and careful interpretation of the predictive results.

Table 1a. Data audit and chronological validation structure

Item	Evidence from dataset	Implication for the article
Panel coverage	21 banks observed annually from 2010 to 2024	Confirms a balanced 15-year panel
Bank-type split	18 conventional banks and 3 fully fledged Islamic banks	Supports separate bank-type analysis
Observations	315 bank-year observations: 270 conventional and 45 Islamic	Clarifies sample size and statistical power
Completeness check	No missing values detected in the model variables	Supports reproducible estimation
Chronological validation	2010-2021 for training and 2022-2024 for testing	Avoids look-ahead bias in prediction
Test observations	54 conventional and 9 Islamic observations in the test period	Explains cautious interpretation of Islamic-bank AI results

The main study variables are measured using the following ratios:

$ROA_{i,t}$ = Net $income_{i,t}$ / Total $assets_{i,t}$ (1)

$NPL_{i,t}$ = non-performing $loans_{i,t}$ / Total $loans_{i,t}$ (2)

$LR_{i,t}$ = Liquid $assets_{i,t}$ / Total $deposits_{i,t}$ (3)

$SER_{i,t}$ = Total $liabilities_{i,t}$ / Total $assets_{i,t}$ (4)

In (1) - (4) i is the bank, and t is the year. $ROA_{i,t}$ is Return on Assets and it is a profitability-based financial-stability indicator. Credit risk is expressed by NPL_i, which stands for Non-Performing Loans ratio. $LR_{i,t}$ is the Liquid Assets-to-Deposits Ratio, which is a measure of liquidity position. SOLVENCY or LEVERAGE Risk: SER_i is the ratio of Liabilities to Assets. The higher the ROA, the more profitable the business; the higher the NPL, the lower the asset quality. LR includes liquidity buffers while SER includes the level of balance-sheet leverage.

Stationarity and integration-order screening

The variables were checked for stationarity prior to estimating the dynamic models. Fisher type panel ADF tests were performed with trend and intercept on bank-level annual series for bank-level variables. The corresponding annual time series were used to screen for GDP growth and inflation. The results show that most variables start to be stationary after taking first difference, but LR is stationary at level. This is addressed by a dynamic error-correction structure, and by being cognizant of the mixed persistence structure as well as conducting robustness checks with Difference GMM.

Table 1b. Stationarity and integration-order screening

Variable	Test form	Level p-value	First-difference p-value	Inference
ROA	Panel ADF-Fisher	0.0624	<0.001	Stationary after differencing at 5%
NPL	Panel ADF-Fisher	0.0949	<0.001	Stationary after differencing at 5%
LR	Panel ADF-Fisher	0.0039	<0.001	Stationary at level
SER	Panel ADF-Fisher	0.4679	<0.001	Stationary after differencing at

				5%
GDP growth	ADF time-series	0.0622	<0.001	Stationary after differencing at 5%
Inflation	ADF time-series	0.6557	<0.001	Stationary after differencing at 5%

4. Methodology

4.1 Explanatory econometric strategy

The article is designed multi-method because this article has two functions, which are explanatory and predictive. The explanatory component starts from descriptive statistics and correlation to find out differences in risk profile among bank types. When variables have long-run equilibrium relationships but, perhaps, short-run deviations from the long run, the dynamic component employs Panel Vector Error Correction Modelling (Panel VECM). To address the persistence of profitability, endogeneity and unobservable bank specific effects, Difference Generalised Method of Moments (Difference GMM) is used as a robustness estimator.

The econometric framework is defined as:

$$Y_{i,t} = [ROA_{i,t}, NPL_{i,t}, LR_{i,t}, SER_{i,t}]' \quad (5)$$

$$\Delta Y_{i,t} = \alpha\beta' Y_{i,t-1} + \sum_{k=1}^{p-1} \Gamma_k \Delta Y_{i,t-k} + \mu_i + \varepsilon_{i,t} \quad (6)$$

$$ROA_{i,t} = \rho ROA_{i,t-1} + \beta_1 NPL_{i,t} + \beta_2 LR_{i,t} + \beta_3 SER_{i,t} + \beta_4 GDP_t + \beta_5 INF_t + \eta_i + u_{i,t} \quad (7)$$

$$\Delta ROA_{i,t} = \rho \Delta ROA_{i,t-1} + \beta_1 \Delta NPL_{i,t} + \beta_2 \Delta LR_{i,t} + \beta_3 \Delta SER_{i,t} + \beta_4 \Delta GDP_t + \beta_5 \Delta INF_t + \Delta u_{i,t} \quad (8)$$

The vector of banking stability and risk variables for banking institution i and year t is denoted by $Y_{i,t}$ in the equations (5)-(8). The symbol Δ denotes the first difference, or the change from one year to the next. GDP_t represents growth of GDP over the past year and INF_t represents inflation over the past year. The Panel VECM specification has the following notation: α, the adjustment coefficient matrix; β, the cointegrating vector or matrix; Γ_k, the short-run dynamic coefficient matrix at lag k; p, the lag length selected; μ_i, the bank-specific effects; and ε_i, the error term.

The dynamics in the short run and the long run are included in this equation. The long-run cointegrating relationship is represented by $\beta' Y_{i,-1}$, and the speed of convergence to the long-run equilibrium after disequilibrium by α. The ROA equation's adjustment coefficient is negative and statistically significant, which means that the equation is moving back toward long-run equilibrium. The analysis is complemented by Difference GMM which employs lagged levels as instruments for differenced endogenous variables, and tests the specification using the AR (1), AR (2), Hansen and Sargan diagnostics.

4.2 Predictive machine-learning and deep-learning strategy

The predictive part considers the forecasting task of ROA as supervised learning problem. The risk indicators and macroeconomic controls are included in the input vector and ROA is included in the target variable. Islamic bank models are estimated separately from conventional bank models and thus one model is not imposed on two banking systems with different operating principles. The final prediction exercise is based on chronological validation: the 2010 to 2021 observations are used for training and the 2022 to 2024 observations are kept for testing. For neural-network models, the predictors are scaled based on information from the training sample, and for tree-based models, the predictors are not scaled.

Random Forest, Multilayer Perceptron (MLP), Hybrid RF-MLP as a combined tree-neural model, LSTM and GRU for temporal dependence. Out-of-sample R^2, Root Mean Squared Error (RMSE), and Mean Absolute Error (MAE) are used to evaluate the performance.

The predictive framework is defined as:

$$X_{i,t} = [NPL_{i,t}, LR_{i,t}, SER_{i,t}, GDP_t, INF_t] \quad (9)$$

$$y_{i,t} = ROA_{i,t} \quad (10)$$

$$\hat{y}_{RF,i,t} = \frac{1}{B}\sum_{b=1}^{B} T_b\left(X_{i,t}\right) \quad (11)$$

$$h = \sigma\left(W_1 X_{i,t} + b_1\right), \quad \hat{y}_{MLP,i,t} = W_2 h + b_2 \quad (12)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \quad (13)$$

$$MAE = \frac{1}{n}\sum_{t=1}^{n}|y_t - \hat{y}_t| \quad (14)$$

$$R^2 = 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2} \quad (15)$$

In Equations (9)-(15), $X_{i,t}$ is the input vector for bank i in year t, and $y_{i,t}$ is the target variable. $\hat{y}_{RF,i,t}$ is the predicted ROA from the Random Forest model, B is the total number of regression trees and $T_b(\cdot)$ is the prediction from the b-th tree. In the MLP equation, h is the hidden-layer representation, $\sigma(\cdot)$ is the activation function, W_1 and W_2 are weight matrices, b_1 and b_2 are bias terms and $\hat{y}_{MLP,i,t}$ is the predicted ROA from the MLP model. For forecast evaluation, y_t is the actual ROA in the test sample, $\hat{y}_t$ is the predicted ROA, $\bar{y}$ is the mean of actual ROA in the test sample and n is the number of test observations. RMSE measures the average size of prediction errors with stronger penalty for large errors, MAE measures the average absolute prediction error and out-of-sample R^2 measures improvement over the mean benchmark.

Because the dataset is annual and the Islamic-bank sample is small, weak forecast scores are treated as diagnostic evidence of model-data mismatch rather than being excluded. This improves transparency by showing when model complexity is not supported by the available banking data. To strengthen reproducibility, the predictive layer was implemented under a restricted and time-aware design. Hyperparameter selection was conducted within the training sample only, and the 2022-2024 observations were kept fully out of sample. Neural-network models used scaled predictors, whereas Random Forest was estimated on the original feature scale because tree-based splits are scale-insensitive.

Table 1c. Predictive-model implementation and validation details

Model/step	Implementation detail	Purpose
Chronological split	2010-2021 training; 2022-2024 testing	Prevents look-ahead bias
Random Forest	Regression forest with bootstrap aggregation; tuning restricted to training data	Nonlinear tabular benchmark
MLP	Feed-forward neural network using scaled predictors and MSE loss	Neural nonlinear benchmark
Hybrid RF-MLP	Random Forest representation followed by MLP prediction	Tests whether tree-based feature learning improves neural prediction

LSTM	Recurrent neural network for longer-memory temporal dependence	Captures persistence in annual banking indicators
GRU	Compact recurrent network with fewer gates than LSTM	Reduces parameter burden in a small panel
Evaluation	Out-of-sample R^2, RMSE and MAE	Compares benchmark improvement and prediction error

Table 1d. Validation and interpretation protocol

Validation issue	Procedure used	Interpretation boundary
Econometric specification	Lag-order selection, cointegration evidence and diagnostic checks are reported before interpreting VECM and GMM results.	Econometric coefficients are treated as explanatory evidence, not direct causal proof beyond model assumptions.
Endogeneity and persistence	Difference GMM is assessed through AR (1), AR (2), Hansen and Sargan diagnostics.	GMM findings are interpreted cautiously where diagnostics are mixed.
Predictive validation	AI models are assessed using out-of-sample R^2, RMSE and MAE under chronological validation.	Negative R^2 values are retained as evidence of limited benchmark improvement.
Islamic-bank sample size	Islamic and conventional banks are estimated separately.	Islamic-bank AI results are treated as exploratory and sample-sensitive.
Policy use	Policy implications are linked to findings supported across descriptive, dynamic and predictive evidence.	AI outputs are positioned as decision-support tools, not automatic decisions.

5. Empirical Results and Interpretation

5.1 Descriptive risk profile

The conventional and Islamic banks descriptive statistics are compared in Table 2. Islamic banks have higher average ROA than conventional banks, and lower NPLs and less volatility of profitability. This means that Islamic banking institutions not only showed greater profitability on average, but also greater stability in terms of profitability. The biggest disparity is in the NPL ratio, where the mean for conventional banks is 12.2149 per cent, while that of Islamic banks is 5.3189 per cent. Generally, liquidity ratios are in similar range and for Islamic banks, SER is lower indicating good solvency ratios. The interpretation is kept on hold due to the limited number of fully operational Islamic banks in the Islamic-bank sample, but it is valid as it portrays the true scenario of Pakistan's dual banking sector.

Table 2. Summary statistics by banking system, 2010-2024

Bank type	Variable	Mean	Std. Dev.	Minimum	Maximum
Conventional	ROA	0.7694	1.5175	-8.9600	4.3600

Conventional	NPL	12.2149	11.6773	1.0400	69.9500
Conventional	LR	8.1778	2.6026	1.1800	16.6000
Conventional	SER	8.0857	7.1707	-12.9300	76.1800
Islamic	ROA	0.9249	0.6342	-0.0500	2.8500
Islamic	NPL	5.3189	4.3494	1.3600	22.8300
Islamic	LR	8.5884	2.4852	5.0900	16.6800
Islamic	SER	7.0642	2.1611	4.6100	15.1600
Both systems	GDP	3.4820	2.23	-1.2700	6.5700
Both systems	Inflation	10.5320	6.93	2.5300	30.7700

ROA, NPL, LR and SER are reported as percentages.

The descriptive results of the data in Table 2 support the hypothesis H4 at the profile level. The Islamic banks have higher profitability and lower credit risk and volatility in profitability. The NPLs and SER of conventional banks are much more spread out, reflecting higher asset quality and balance sheet stress. The table does not imply causality but only suggests the need to pursue the dynamic VECM and GMM analysis that is to follow.

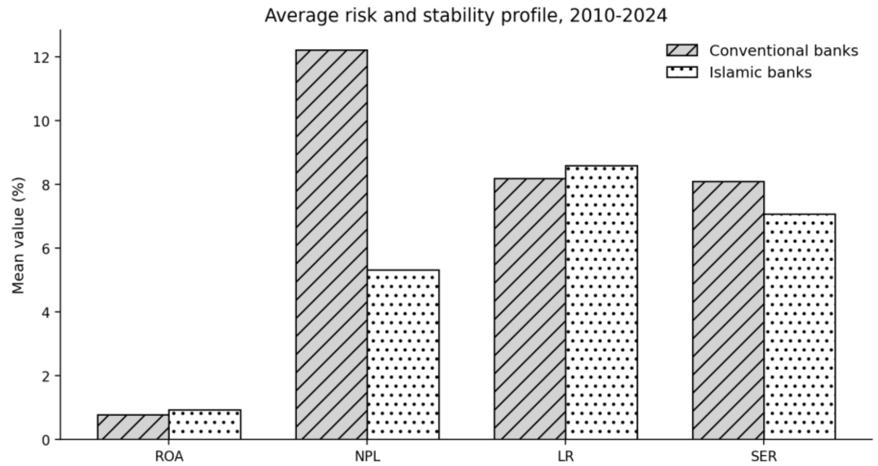

Figure 2. Average risk and stability profile of conventional and Islamic banks.

5.2 Correlation evidence

The correlation matrix confirms the key role of credit risk. The correlation between ROA and NPL is negative and -0.5533 for conventional banks. In Islamic banks, the correlation is -0.3888. These ratings suggest that the deterioration of the quality of the assets is always linked with less profitability. The negative association is more robust in the conventional sample, which corroborates the hypothesis that the credit-risk pressure is higher in conventional banks. Liquidity is less highly related to ROA in both systems, suggesting that liquidity is not the primary factor for stability. The positive simple correlation of inflation and ROA of Islamic banks does not mean that inflation is good; it merely indicates an unconditional association that needs to be tested in a controlled dynamic model.

Table 3. Selected correlations with ROA

Bank type	Corr(ROA,NPL)	Corr(ROA,LR)	Corr(ROA,SER)	Corr(ROA,GDP)	Corr(ROA,Inflation)
Conventional	-0.5533	-0.0224	0.2393	-0.0238	0.0449
Islamic	-0.3888	-0.1477	-0.3113	-0.2799	0.5466

The main profitability channel is the credit risk, as also indicated in Table 3. Conventional banks have a stronger negative correlation between ROA and NPL, signifying that they are more credit-cycle sensitive. The low level of liquidity correlation suggests that liquidity buffers are important and that their influence could be positive or negative. The opposite SER correlations between systems support separate modelling over a pooled-only modelling approach.

5.3 Lag-order and cointegration results

Two lags are provided for conventional and Islamic bank selection for lag-order selection. Johansen cointegration tests show that there is a long-run equilibrium relationship in both systems. The trace statistic for conventional banks suggests two cointegrating relationships at 5 per cent, and the maximum-eigenvalue test suggests at least one. Trace and maximum-eigenvalue tests both indicate that there are two cointegrating relationships for Islamic banks. These findings validate the VECM specification as well as corroborate the notion that the liquidity, credit, solvency and profitability are factors of a long-run equilibrium structure.

Table 4. Lag-order and cointegration summary

Test item	Conventional banks	Islamic banks	Interpretation
Selected lag order	2	2	Both systems require dynamic modelling with two lags
Trace test	Two cointegrating equations	Two cointegrating equations	Long-run equilibrium exists in both systems
Maximum eigenvalue test	At least one cointegrating equation	Two cointegrating equations	Islamic-bank equilibrium evidence is especially strong
Model implication	Use Panel VECM	Use Panel VECM	Short-run dynamics and long-run adjustment should be modelled jointly

The dynamic modelling strategy given in the article is validated in Table 4. A VECM specification would not be as defensible without the presence of cointegration. Both systems exhibit LR equilibrium, so the adjustment coefficients can be taken as indications of the way in which profitability returns to the LR equilibrium state following risk shocks.

Below, the full evidence of the lag-selection are presented to make the dynamic specification clear. Since there are 15 time points per year, the maximum candidate lag was limited to two to prevent over-parameterization. Conventional and Islamic bank systems choose lag two in AIC, BIC, HQIC and FPE.

***Table 4a**. Detailed lag-order selection results*

Bank type	Lag	AIC	BIC	HQIC	FPE	Decision
Conventional	0	-0.4891	-0.3152	-0.5248	0.6139	Not selected
Conventional	1	-2.6530	-1.7840	-2.8320	0.0832	Not selected
Conventional	2	-4.6907	-3.1262	-5.0122	0.0330	Selected
Islamic	0	-1.2740	-1.1000	-1.3090	0.2801	Not selected
Islamic	1	-3.2810	-2.4120	-3.4600	0.0444	Not selected
Islamic	2	-14.5198	-12.9554	-14.8414	1.779e-06	Selected

The Johansen statistics show in detail that a VECM is appropriate. The trace statistic can be used to test for 2 cointegrating relations for conventional banks, and the maximum-eigenvalue statistic for at least one. Islamic banks have two cointegrating relations in both trace and maximum-eigenvalue statistics. These findings support the modelling of long-run adjustment as opposed to merely static regression.

***Table 4b**. Johansen cointegration test details*

Bank type	Null hypothesis	Trace statistic	5% trace critical	Max-eigen statistic	5% max-eigen critical	Decision
Conventional	r = 0	79.9113	47.8545	49.9253	27.5858	Reject no cointegration
Conventional	r ≤ 1	29.9861	29.7961	17.8198	21.1314	Trace supports second relation; max-eigen does not
Conventional	r ≤ 2	12.1663	15.4943	11.1419	14.2639	Do not reject
Islamic	r = 0	203.7622	47.8545	155.3447	27.5858	Reject no cointegration
Islamic	r ≤ 1	48.4176	29.7961	36.9779	21.1314	Reject; second relation supported
Islamic	r ≤ 2	11.4396	15.4943	11.3528	14.2639	Do not reject

5.4 Panel VECM results

The main VECM results for ROA equation are reported in Table 5. At conventional banks, the value of ECT1 is negative and substantial, suggesting that disequilibrium has a strong effect on profitability. ECT2 is positive, indicating that the second long-run relation is an alternative path. The coefficient for lagged NPL is negative and statistically significant, as they are expected to make profits less. Both ECT1 and ECT2 have negative and significant values in the ROA equation for Islamic banks. They are smaller, indicating a slower but more stable adjustment trajectory. The results do not mean that Islamic banks are risk free, but rather it shows that Islamic banks have a different adjustment structure, and that profitability is not as sharply adjusted following shocks.

Table 5. Key VECM results for the ROA equation

Bank type	ECT1(-1)	ECT2(-1)	D(ROA(-1))	D(NPL(-1))	D(LR(-1))	D(SER(-1))
Conventional	-2.5304 [-7.2998]	0.7056 [7.4877]	0.5364 [2.9097]	-0.6151 [-6.2604]	0.4119 [5.1253]	-0.0189 [-0.2823]
Islamic	-0.3011 [-2.6422]	-0.1274 [-2.6934]	0.1456 [0.4469]	0.0782 [2.2571]	-0.1040 [-1.4016]	0.0895 [0.5832]

t-statistics are shown in brackets. The table reports the ROA equation only to focus on financial-stability adjustment.

As indicated in Table 5, the VECM structure of conventional banks has a larger and quicker profitability correction, alongside a sudden negative effect on short-term NPLs. This suggests that the profitability of conventional banks is more sensitive to credit-quality shocks, and reinforces H1. The adjustment coefficients of Islamic banks are in interesting range, but smaller, which lend support to H4 via a somewhat more gradual risk-adjustment mechanism. Care needs must be taken with the Islamic-bank NPL coefficient, due to the smaller sample.

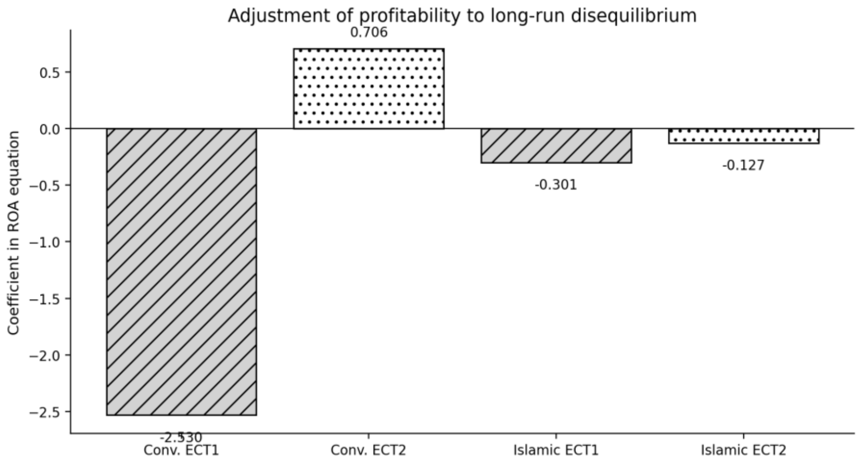

Figure 3. ROA adjustment coefficients from the VECM.

5.5 Difference GMM robustness results

The difficulty of endogeneity, dynamic persistence and reverse causality is addressed by Difference GMM. The results of the diagnostic test support the chosen instrument set. The AR (2) p values of both conventional and Islamic banks are greater than the conventional level of significance, which means there is no evidence of second-order serial correlation in the differenced residua. The null hypothesis of instrument validity is not rejected by Hansen J tests. The Sargan test is not accepted by conventional banks, but it is accepted by Islamic banks; the sensitivity of the Sargan statistic to heteroskedasticity makes the outcome of the Hansen test more reliable in this context.

Table 6. GMM diagnostics and coefficient summary

Item	Conventional banks	Islamic banks	Decision/meaning
AR(1) p-value	0.0000	0.1637	First-order correlation expected in differenced errors
AR(2) p-value	0.3316	0.4229	No evidence of second-order serial correlation
Hansen J p-value	0.6786	0.4332	Do not reject instrument validity
Sargan p-value	0.0010	0.9556	Rejects only in the conventional sample; interpreted cautiously
Significant coefficients	None at 5 per cent	Inflation, lagged ROA, SER	Islamic-bank profitability is more sensitive to persistence and solvency

The results presented in Table 6 makes the use of GMM as a robustness check plausible. The lack of AR (2) is especially relevant as second-order serial correlation would make the instruments less valid. The results indicate that the Islamic-bank model is clear of significant coefficients and conventional model is weak in terms of short-run GMM effects after the endogeneity correction.

The Difference GMM specification is not reported as the only ground for inference, but as a robustness layer. The instrument strategy relies on the assumption of no AR (2) serial correlation and non-rejection of the Hansen validity test for the endogenous variables under differencing. This reporting helps to minimize the possibility of misinterpreting dynamic coefficients as causal effects without proper diagnostic support.

Table 6a. Difference GMM implementation and diagnostic support

GMM element	Conventional banks	Islamic banks	Interpretation
Groups	18 banks	3 banks	Reflects the actual structure of the sample
Observations	270 bank-year observations	45 bank-year observations	Islamic estimates require cautious interpretation
Estimator role	Robustness estimator	Robustness estimator	Addresses persistence and potential endogeneity
Instrument logic	Lagged levels for differenced endogenous variables	Lagged levels for differenced endogenous variables	Internal instruments reduce simultaneity concerns
Serial correlation check	AR(2) p = 0.3316	AR(2) p = 0.4229	No evidence of second-order serial correlation
Instrument validity check	Hansen p = 0.6786	Hansen p = 0.4332	Do not reject instrument validity
Caution	Sargan rejects	Sargan does not reject	Hansen result is emphasised because Sargan is sensitive to heteroskedasticity

Table 7. Difference GMM coefficient estimates for Delta ROA

Term	Conventional coefficient	Conventional p-value	Islamic coefficient	Islamic p-value
D(GDP)	-0.0304	0.2405	0.0207	0.2908
D(Inflation)	0.0091	0.2853	0.0502	0.0000
D(ROA(-1))	0.2167	0.3887	0.8459	0.0050
D(NPL)	-0.0355	0.6733	0.0125	0.5142
D(LR)	-0.0171	0.8529	0.0035	0.9183
D(SER)	0.1796	0.2683	0.1822	0.0390

In the conventional model, as well as differenced and instrumented, no variable is significant at the 5 per cent level as displayed in Table 7. This implies that when endogeneity and persistence are controlled for, it becomes difficult to explain the short-run variation in ROA. For Islamic bank, inflation and lagged ROA and SER are significant, which shows profitability persistence and higher sensitivity to macro-economic and solvency related factors. The positive SER coefficient should not be interpreted as an indicator of the positive effect of higher liabilities on the performance of Islamic banks, but could simply be a result of the financing and balance-sheet structure of Islamic banks in the sample period.

5.6 Machine-learning and deep-learning forecasting results

Predictive performance is reported in Table 8. Conventional-bank ML results indicate that despite the negative out-of-sample R^2, the strongest machine-learning model is Random Forest of the three models considered (RF, MLP, Hybrid RF-MLP). This means that the selected annual prediction setting is not met by the model. The deep-learning results are better for conventional banks: The out-of-sample R^2, RMSE, and MAE values for the conventional banks are 0.4033, 0.8492, and 0.5122, respectively, of which the former is best. Among the deep-learning layer, LSTM outperforms GRU for Islamic banks and Hybrid RF-MLP has the lowest machine-learning error. The negative R^2 values, however, suggest that complex AI models should be used with caution when the sample size for the Islamic bank is small.

Table 8. Predictive performance of ML and DL models

Bank type	Model	Out-of-sample R^2	RMSE	MAE	Interpretation
Conventional	Random Forest	-0.1143	1.1604	0.7949	Best ML model for conventional banks
Conventional	MLP	-2.2176	1.9718	1.6949	Weak due to overfitting/limited annual sample
Conventional	Hybrid RF-MLP	-5.2152	2.7405	1.8883	Weakest conventional ML result
Conventional	LSTM	0.2366	0.9605	0.6204	Useful temporal model
Conventional	GRU	0.4033	0.8492	0.5122	Best overall conventional predictive model
Islamic	Random Forest	-2.0089	1.0966	0.9736	Weak but interpretable benchmark
Islamic	MLP	-5.1029	1.5618	1.4707	Weak neural model
Islamic	Hybrid RF-MLP	-1.3536	0.9699	0.8490	Best Islamic ML model by error
Islamic	LSTM	-0.4676	0.7659	0.6317	Best Islamic DL model by error
Islamic	GRU	-0.7348	0.8327	0.7474	Weaker than LSTM for Islamic banks

Table 8 helps to support H5 by demonstrating that the econometric and AI models provide complementary evidence, and that AI is not necessarily better. The forecast properties of GRU seems to be of relevance for conventional banks, while for Islamic banks, the error minimisation and parsimony of the model may be more important than the out-of-sample R^2. The negative out of sample R^2 values are kept since they demonstrate that deep learning models require a large sample size, a higher number of observations or more complex features to be used as supervisory early-warning tools for Islamic banks.

In addition to that, a simple linear regression model was estimated with the same predictors and with the same chronological train-test split as above as check. The benchmark proves AI isn't always better than simpler models. GRU is still the top conventional-bank forecasting model, and the Islamic-bank results are sensitive to the samples and should be seen as exploratory.

Table 8a. Linear and mean benchmark forecasting performance

Bank type	Benchmark	Out-of-	RMSE	MAE	Interpretation

	model	sample R^2			
Conventional	Mean benchmark	-0.0743	1.1394	0.7356	Competitive with ML
Conventional	Linear regression	-0.2267	1.2175	0.7563	Below GRU
Islamic	Mean benchmark	-2.3487	1.1569	0.9689	Weak test mean
Islamic	Linear regression	-0.4511	0.7616	0.6508	Close to LSTM

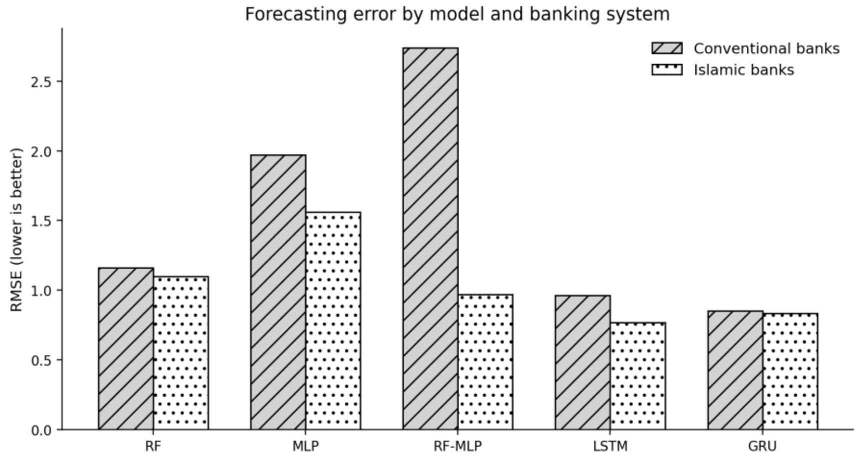

Figure 4. Forecasting error by model and banking system.

5.7 Hypothesis-wise synthesis

Table 9. Hypothesis-wise summary of empirical support

Hypothesis	Expected relationship	Evidence used	Decision
H1	Credit risk reduces stability and differs by bank type	Descriptive NPL gap; ROA-NPL correlations; VECM NPL effect	Supported, especially for conventional banks
H2	Liquidity affects stability differently across systems	Descriptive LR; correlation; VECM LR coefficient	Partially supported; liquidity effects are weaker and context-dependent
H3	Solvency affects long-run equilibrium and adjustment	Cointegration; VECM; GMM SER significance in Islamic banks	Supported with stronger evidence in dynamic models
H4	Islamic and conventional banks differ in resilience	ROA/NPL/SER profiles; VECM adjustment speed	Supported, with cautious interpretation due to smaller Islamic sample
H5	Econometric and AI evidence is complementary	GMM/VECM plus ML/DL forecast performance	Supported; GRU useful for conventional banks, AI weaker for small Islamic panels

The main contribution of the article is clarified by a hypothesis-wise synthesis in table 9. The best evidence is for credit risk and bank-type differences. Liquidity effects are less straightforward and more conditional, as they may be either positive or negative depending on the specifics. The evidence that is predictive supports the idea that AI should complement, not supplant, econometric explanation in supervisory situations.

6. Discussion

The overall evidence indicates three principal interpretations. First, the greatest risk channel in Pakistan's banking system is credit risk. This is corroborated by descriptive statistics, correlations and VECM dynamics. Conventional banks have significantly higher NPLs while the correlation of ROA to NPLs is negative. The result is in line with literature which shows that asset quality is one of the factors that explain bank fragility (Ghenimi et al., 2017; Abdelaziz et al., 2022; Naili & Lahrichi, 2022).

Second, on average, Islamic banks do not seem to be as fragile, although they are not aggressive in being resilient. Islamic banking institutions exhibit higher average profitability, lower credit risk and lower volatility in profitability. The correction coefficients for their VECM are smaller, but not negligible, indicating a less rapid correction to equilibrium. This is in line with the opinion of Islamic banking structures having less risk exposure, due to the asset-backing and contractual discipline, but still constrained by liquidity and macroeconomic factors. The results of this study corroborate and extend the results of Beck et al. (2013), Safiullah (2021), Ghenimi et al. (2024) and Sendi et al. (2024).

Thirdly, the predictive layer reshapes the definition of AI in finance. The notion that deep learning is "always best" is not supported by the evidence. GRU is effective for conventional banks but still has limits in Islamic banks due to smaller sample size. Hence, the article provides an optimistic yet pragmatic reminder: AI models can help with early warning, but they need to be adapted to the volume of data, its frequency and type of banks, as well as the complexity of the bank model. In banking-risk research, a negative out-of-sample R^2 is not just a bad performance—it is a red flag to the supervisors that the model should not be put into use for that segment without additional data, additional features, more frequent observations.

The integrated econometric-AI design enhances the policy relevance. The VECM and GMM models provide information about the composition of risk channels and whether they can provide usable forecasts; the former yields information about the composition of VECM, while the latter yields information about the composition of GMM. Without prediction, explanation can only recognize historical patterns, which is not very effective as early warning. Explanation-free prediction might result in a black-box score that may not be easy to justify in a supervisory context. So, the contribution of the article is not model proliferation but model integration.

7. Contributions and Implications

7.1 Theoretical contribution

The article adds to financial stability theory by demonstrating how the risk resilience can be characterized as a combination of asset quality, liquidity capacity and solvency position. It supports the dual banking theory in that Islamic and conventional banks are not just different in terms of their average profitability but also in terms of their adjustment speed, credit risk exposure, and in terms of the suitability of their credit risk predictive models. This takes the Islamic-conventional comparison out of the descriptive benchmarking arena. The results also clarify the notion that Islamic banks inherently are more stable. Islamic banks look better in terms of average profitability and credit risk, whereas their predictive modelling results are not as good due to the smaller sample size. There is hence a connection between institutional resilience and statistical predictability, but they do not mean the same.

7.2 Methodological contribution

Methodologically, the article illustrates the application of the combination of dynamic econometrics and AI forecasting without mixing up their roles in an integrated paper on banking risk. Panel VECM detects long-run equilibrium and short-run adjustment, Difference GMM tackles with dynamic endogeneity, and the Random Forest, MLP, Hybrid RF-MLP, LSTM and GRU test the predictive usefulness. The framework is particularly applicable to emerging-market banking systems, where relationships are fluid, and samples are sometimes limited.

7.3 Policy and managerial implications

The credit-risk surveillance is the initial line of defence for the State Bank of Pakistan as NPLs are the strong recurring risk channel. Liquidity, credit and solvency indicators must be seen as a system, and not as independent ratios that are reported. Islamic banks need more liquidity markets and better risk-sharing governance that is compliant with Islamic principles to ensure the advantages that have been observed in terms of stability are continued. Conventional banks need to build up early warning systems for deterioration in asset quality, particularly in inflationary or times of macroeconomic stress. Just as with other tools, one should use AI tools only after validating that they work with the data, and for Islamic-bank forecasting, larger samples and more frequent data or other explanatory factors are needed.

7.4 Policy-action matrix

Table 10 then assigns each risk signal to a regulatory and bank-level response; to provide a translation of the empirical evidence into an action the bank can take to add to its supervisory contribution. The discussion on the practical contribution in this section illustrates the transformation of the integrated econometric-AI framework into monitoring rules, not just a statistical exercise.

Table 10. Policy and managerial action matrix

Risk signal	Main empirical evidence	Supervisory implication	Bank-level implication
Credit risk / NPLs	NPLs show the strongest negative association with ROA, especially for conventional banks.	Prioritise asset-quality surveillance, provisioning discipline and stress tests around loan-quality deterioration.	Strengthen credit screening, sector concentration limits and early-warning triggers for borrower distress.
Liquidity position	Liquidity effects are weaker and more bank-type dependent than credit-risk effects.	Monitor liquidity jointly with credit and solvency, rather than treating LR as an isolated compliance ratio.	Maintain liquidity buffers while avoiding excessive low-yield asset holdings that can reduce profitability.
Solvency risk	Cointegration and GMM evidence show that solvency is part of the long-run stability mechanism.	Integrate leverage and capital-buffer indicators into macroprudential resilience dashboards.	Link capital planning with stress testing, loan growth and balance-sheet expansion.
Islamic-bank resilience	Islamic banks show lower NPLs and lower ROA volatility, but slower adjustment and smaller sample size.	Develop deeper Shariah-compliant liquidity instruments and separate supervisory calibration for Islamic banks.	Improve Shariah-compliant liquidity planning and risk-sharing governance while expanding data infrastructure.
AI forecasting	GRU performs best for conventional banks,	Use AI as a validated early-warning	Deploy models only after bank-type-specific

	while Islamic-bank AI results are constrained by sample size.	supplement, not as a black-box replacement for econometric evidence.	validation, back-testing and explainability checks.

Table 10 interprets the empirical results to actions for supervisors and managers. The most important regulatory interest is credit-risk monitoring, though the interest in liquidity and solvency should be considered together. The AI outputs should not be used for automatic warnings, but for model assisted warnings.

8. Robustness, Validity Checks, Limitations and Future Research

The results of the econometrics are supplemented by the Difference GMM diagnostics. There is no second order serial correlation as indicated by the AR (2) test and instrument validity is not rejected by the Hansen J test. Lag-order selection and cointegration tests support the VECM result. Several constraints have to be noted, however. First, ROA is a useful, but limited indicator of stability. Other ratios which might be worth considering for future research include Z-scores, capital adequacy ratios, liquidity coverage ratios, non-performing financing ratios or market-based stability indicators.

Second, the smaller size of the Islamic-bank sample is due to the fact that in Pakistan, there are fewer fully fledged Islamic banks. This reduces the statistical power of the AI models and calls for a careful interpretation. Thirdly, the data are annual. For deep learning, annual data are less suitable because they are noisy, whereas for econometrics they are more suitable. If quarterly or monthly data is available, future research could use this data. Fourth, some financial and macroeconomic variables are limited in the predictive models. Other metrics that could be explored in future studies are governance indicators, bank size, capital adequacy, ESG performance, Shariah governance, digitalisation, deposit concentration, and market sentiment. Fifth, future research needs could be conducted on the other dual banking systems of the world like Malaysia, Indonesia, Turkey or GCC countries to see the country specificity or generalization of the results.

8.1 Robustness and validity-check synthesis

Table 11 summarizes the credibility protection applied to the article. The aim of presentation is not to exaggerate the findings but to demonstrate that the findings are backed by a multi-layered process of evidence: descriptive comparison, dynamic cointegration, endogeneity correction, prediction validation, and explicit consideration of model specification, etc.

Table 11. Robustness and validity-check synthesis

Validity concern	Check used in the manuscript	Interpretation	Effect on interpretation strength
Dynamic specification	Lag-order selection indicates two lags for both bank groups.	The banking variables require dynamic modelling rather than a static regression-only design.	Strengthens the use of Panel VECM.
Long-run relationship	Johansen cointegration confirms long-run equilibrium relationships in both systems.	ROA, NPL, LR and SER move as part of a connected stability structure.	Supports the integrated risk framework.
Endogeneity	Difference GMM uses	AR(2) and Hansen results	Improves credibility of

and persistence	internal instruments and reports AR(2), Hansen and Sargan diagnostics.	support the selected instrument strategy, with Sargan interpreted cautiously.	dynamic profitability estimates.
Prediction and explanation	ML/DL models are evaluated with R-squared, RMSE and MAE.	Negative out-of-sample R^2 values are reported rather than hidden, showing when AI is unsuitable.	Strengthens transparency and avoids inflated AI claims.
Small Islamic-bank sample	Islamic and conventional banks are estimated separately and limitations are explicitly stated.	The Islamic sample reflects the actual population of fully fledged Islamic banks, but results remain sample-sensitive.	Supports cautious and defensible interpretation.
Policy relevance	Results are translated into bank-type-specific supervisory implications.	Findings are actionable for risk committees, regulators and Islamic finance supervisors.	Strengthens practical contribution.

From Table 11, it is observed that it is not a model dependent article. Descriptive, correlation, VECM and GMM evidence in agreement yields the strongest claims. AI-based claims are typically limited to predictive statements, rather than causal claims, which enhances the credibility of the proposed interpretation.

8.2 Validity boundaries and interpretation calibration

Table 12 below outlines the major issues with validity and how the manuscript addresses those issues without going too far. This enhances interpretive discipline, so that every conclusion is related to its level of support and its limits.

Table 12. Validity boundaries and interpretation checklist

Validity concern	How the manuscript addresses it	Remaining claim boundary
Small Islamic-bank sample	The paper treats the Islamic sample as the observed population of fully fledged Islamic banks, estimates models separately and interprets AI results cautiously.	Findings are strongest for Pakistan and should not be universalized without multi-country testing.
Use of ROA as stability proxy	ROA is justified as an earnings-based resilience indicator and linked with risk channels, while limitations note the value of Z-score, CAR and ROE extensions.	ROA captures profitability resilience, not all dimensions of systemic stability.
Negative AI R-squared values	Negative values are disclosed and interpreted as model-data mismatch rather than hidden. RMSE and MAE are also used for error comparison.	AI findings are predictive diagnostics, not causal results.
Potential endogeneity	Difference GMM is used as a robustness estimator with AR (2), Hansen and Sargan diagnostics.	Causal language is avoided unless supported by the model structure and diagnostics.

Too many models	The manuscript assigns different roles: VECM for equilibrium, GMM for dynamic robustness, ML/DL for predictive early-warning comparison.	The article claims model complementarity rather than universal model superiority.
Policy relevance	Results are converted into a bank-type-specific action matrix for credit risk, liquidity, solvency and AI deployment.	Policy implications remain risk-management guidance rather than regulatory prescription.

The manuscript's interpretation boundaries are made explicit. The central conclusions still hold up as the argument recognizes limitations and the most compelling claims are linked to the triangulated evidence, not a single model.

9. Conclusion

In this article, an integrated econometric and AI approach to the analysis of risk-resilience in the dual banking system was developed for Pakistan. It analyzed the impact of liquidity risk, credit risk and solvency risk on financial stability in Islamic and conventional banks, based on annual data of 21 banks during the period 2010-2024. The descriptive evidence indicates that Islamic banks have higher average ROA, lower NPLs, lower volatility of profitability, and higher solvency profiles. The correlation and VECM evidence result indicate that the channel of credit risk stability is the most significant channel of stability, particularly for conventional banks.

The VECM estimates long run equilibrium relations in both banking systems are confirmed by the results. Conventional banks tend to adjust more aggressively following disequilibrium while Islamic banks adjust more slowly and steadily. The results of Difference GMM indicates that the significant variables for Islamic banks are inflation and the lagged profitability and solvency related variables, whereas for conventional banks, the significance of the variables is lower after endogeneity correction. The outcome of the AI models indicates that the GRU model is the most superior model for conventional-bank forecasting, while the LSTM and Hybrid RF-MLP models are relatively better by the error metric for the Islamic bank forecasting. Simpler models continue to be competitive in small annual panels, especially for Islamic banks, according to the additional linear benchmark. Multiple negative out-of-sample R^2 values indicate that complex AI models should not be applied without comparison to benchmarks or increased sample size or frequency of observation.

The main idea is that it is not possible to measure the risk resilience of a dual banking system with only one ratio, or with only one model. An improved supervisory stance should include: clarification of risk channels, a test of dynamic adjustment, and an assessment of the predictive usefulness and the explicitness of model risk governance. AI forecasting is useful for regulators and bank managers if it can be verified by bank type, explained with supporting evidence and understood within the parameters of the data.

10. Declarations

Data availability: Bank level annual financial statements and publically available macro-economic indicators are used in the study. Processed panel used for estimation may be made available by the corresponding author on reasonable request with citation to the original public sources.

Funding: The authors did not receive financial support for the research, authorship or publication of this article.

Conflict of interest: The authors have no conflicts of interest to disclose.

Ethical statement: Only secondary bank-level and macroeconomic data are used for the study. Humans are not involved in the process; no interviews and no private personal information is involved.

References

Abbas, F., Iqbal, S. and Aziz, B. (2019), 'The impact of bank capital, bank liquidity and credit risk on profitability in post-crisis period: A comparative study of the US and Asia', Cogent Economics and Finance, Vol. 7 No. 1. https://doi.org/10.1080/23322039.2019.1605683

Abdelaziz, H., Rim, B. and Helmi, H. (2022), 'The interactional relationships between credit risk, liquidity risk and bank profitability in MENA region', Global Business Review, Vol. 23 No. 3, pp. 561-583. https://doi.org/10.1177/0972150919879304

Abdeljawad, I., Rashid, M., Alia, M.A., Qushtom, R., Irshaid, M. and Sahyouni, A. (2024), 'Cushion hypothesis and credit risk: Islamic versus conventional banks from the MENA region', PLoS ONE, Vol. 19 No. 7. https://doi.org/10.1371/journal.pone.0306901

Akram, H. and Hushmat, A. (2024), 'Bank liquidity creation and solvency risk with moderating role of loan concentration: A comparative study of Islamic and conventional banks in Pakistan and Malaysia', Risk Management, Vol. 26 No. 4.. https://doi.org/10.1057/s41283-024-00154-4

Al-Husainy, N.H.M. and Jadah, H.M. (2021), 'The effect of liquidity risk and credit risk on the bank performance: Empirical evidence from Iraq', IRASD Journal of Economics, Vol. 3 No. 1, pp. 58-67. https://doi.org/10.52131/joe.2021.0301.0025

Aldasoro, I., Cho, C.H. and Park, K. (2022), 'Bank solvency risk and funding cost interactions: Evidence from Korea', Journal of Banking and Finance, Vol. 134. https://doi.org/10.1016/j.jbankfin.2021.106348

Barongo, R.I. and Tibangayuka Mbelwa, J. (2024), 'Using machine learning for detecting liquidity risk in banks', Machine Learning with Applications, Vol. 15. https://doi.org/10.24433/CO.5061127.v1

Basel Committee on Banking Supervision (2011), Core Principles for Effective Banking Supervision, Bank for International Settlements, Basel.

Bawa, J.K. and Basu, S. (2020), 'Restructuring assets reform, 2013: Impact of operational ability, liquidity, bank capital, profitability and capital on bank credit risk', IIMB Management Review, Vol. 32 No. 3, pp. 267-279. https://doi.org/10.1016/j.iimb.2019.10.009

Beck, T., Demirgüç-Kunt, A. and Merrouche, O. (2013), 'Islamic vs. conventional banking: Business model, efficiency and stability', Journal of Banking and Finance, Vol. 37 No. 2, pp. 433-447. https://doi.org/10.1016/j.jbankfin.2012.09.016

Ben Lahouel, B., Taleb, L., Ben Zaied, Y. and Managi, S. (2024), 'Financial stability, liquidity risk and income diversification: Evidence from European banks using the CAMELS-DEA approach', Annals of Operations Research, Vol. 334, pp. 391-422. https://doi.org/10.1007/s10479-022-04805-1

Ben-Ahmed, K., Kasraoui, N. and Soulami, A. (2023), 'Exploring the interplay of credit and liquidity risks: Impacts on banks' profitability', International Journal of Advanced and Applied Sciences, Vol. 10 No. 8, pp. 64-70. https://doi.org/10.21833/ijaas.2023.08.007

Breiman, L. (2001), 'Random forests', Machine Learning, Vol. 45 No. 1, pp. 5-32.

Brunnermeier, M.K. (2009), 'Deciphering the liquidity and credit crunch 2007-2008', Journal of Economic Perspectives, Vol. 23 No. 1, pp. 77-100.

Cho, K., Van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H. and Bengio, Y. (2014), 'Learning phrase representations using RNN encoder-decoder for statistical machine translation', Proceedings of EMNLP 2014, pp. 1724-1734.

Cont, R., Kotlicki, A. and Valderrama, L. (2020), 'Liquidity at risk: Joint stress testing of solvency and liquidity', Journal of Banking and Finance, Vol. 118. https://doi.org/10.1016/j.jbankfin.2020.105871
de Bandt, O., Lecarpentier, S. and Pouvelle, C. (2021), 'Determinants of banks' liquidity: A French perspective on interactions between market and regulatory requirements', Journal of Banking and Finance, Vol. 124. https://doi.org/10.1016/j.jbankfin.2020.106032
de Jesus, D.P. and Besarria, C. da N. (2023), 'Machine learning and sentiment analysis: Projecting bank insolvency risk', Research in Economics, Vol. 77 No. 2, pp. 226-238. https://doi.org/10.1016/j.rie.2023.03.001
Demirgüç-Kunt, A., Pedraza, A. and Ruiz-Ortega, C. (2021), 'Banking sector performance during the COVID-19 crisis', Journal of Banking and Finance, Vol. 133. https://doi.org/10.1016/j.jbankfin.2021.106305
Dibooglu, S., Cevik, E.I. and Al Tamimi, H.A.H. (2022), 'Credit default risk in Islamic and conventional banks: Evidence from a GARCH option pricing model', Economic Analysis and Policy, Vol. 75, pp. 396-411. https://doi.org/10.1016/j.eap.2022.06.006
Djebali, N. and Zaghdoudi, K. (2020), 'Threshold effects of liquidity risk and credit risk on bank stability in the MENA region', Journal of Policy Modelling, Vol. 42 No. 5, pp. 1049-1063. https://doi.org/10.1016/j.jpolmod.2020.01.013
El-Chaarani, H., Ismail, T.H., El-Abiad, Z. and El-Deeb, M.S. (2024), 'The impact of COVID-19 on financial structure and performance of Islamic banks: A comparative study with conventional banks in the GCC countries', Journal of Economic and Administrative Sciences, Vol. 40 No. 4, pp. 769-797. https://doi.org/10.1108/JEAS-07-2021-0138
Gafrej, O. and Boujelbene, M. (2022), 'The impact of performance, liquidity and credit risks on banking diversification in a context of financial stress', International Journal of Islamic and Middle Eastern Finance and Management, Vol. 15 No. 1, pp. 66-82. https://doi.org/10.1108/IMEFM-09-2020-0488
Gafsi, N. (2025), 'Machine learning approaches to credit risk: Comparative evidence from participation and conventional banks in the UK', Journal of Risk and Financial Management, Vol. 18 No. 7.. https://doi.org/10.3390/jrfm18070345
Ghenimi, A., Chaibi, H. and Omri, M.A. (2024), 'Risk and performance of Islamic and conventional banks under COVID-19 pandemic: Evidence from MENA region', Arab Gulf Journal of Scientific Research, Vol. 42 No. 4, pp. 1788-1804. https://doi.org/10.1108/AGJSR-03-2023-0098
Ghenimi, A., Chaibi, H. and Omri, M.A.B. (2017), 'The effects of liquidity risk and credit risk on bank stability: Evidence from the MENA region', Borsa Istanbul Review, Vol. 17 No. 4, pp. 238-248. https://doi.org/10.1016/j.bir.2017.05.002
Ghenimi, A., Chaibi, H. and Omri, M.A.B. (2021), 'Liquidity risk determinants: Islamic vs conventional banks', International Journal of Law and Management, Vol. 63 No. 1, pp. 65-95. https://doi.org/10.1108/IJLMA-03-2018-0060
Ghouse, G., Ejaz, N., Bhatti, M.I. and Aslam, A. (2022), 'Performance of Islamic vs conventional banks in OIC countries: Resilience and recovery during Covid-19', Borsa Istanbul Review, Vol. 22, pp. S60-S78. https://doi.org/10.1016/j.bir.2022.11.020
Hassan, M.K., Khan, A. and Paltrinieri, A. (2019), 'Liquidity risk, credit risk and stability in Islamic and conventional banks', Research in International Business and Finance, Vol. 48, pp. 17-31. https://doi.org/10.1016/j.ribaf.2018.10.006
Hochreiter, S. and Schmidhuber, J. (1997), 'Long short-term memory', Neural Computation, Vol. 9 No. 8, pp. 1735-1780.

IFSB (2025), Islamic Financial Services Industry Stability Report 2025, Islamic Financial Services Board.

Imbierowicz, B. and Rauch, C. (2014), 'The relationship between liquidity risk and credit risk in banks', Journal of Banking and Finance, Vol. 40, pp. 242-256. https://doi.org/10.1016/j.jbankfin.2013.11.030

Kabir, M.N., Worthington, A. and Gupta, R. (2015), 'Comparative credit risk in Islamic and conventional bank', Pacific Basin Finance Journal, Vol. 34, pp. 327-353. https://doi.org/10.1016/j.pacfin.2015.06.001

Levine, R. (2004), 'Finance and growth: Theory and evidence', NBER Working Paper No. 10766.

Naili, M. and Lahrichi, Y. (2022), 'Banks' credit risk, systematic determinants and specific factors: Recent evidence from emerging markets', Heliyon, Vol. 8 No. 2. https://doi.org/10.1016/j.heliyon.2022.e08960

Oino, I. (2021), 'Bank solvency: The role of credit and liquidity risks, regulatory capital and economic stability', Banks and Bank Systems, Vol. 16 No. 4, pp. 84-100. https://doi.org/10.21511/bbs.16(4).2021.08

Petropoulos, A., Siakoulis, V., Stavroulakis, E. and Vlachogiannakis, N.E. (2020), 'Predicting bank insolvencies using machine learning techniques', International Journal of Forecasting, Vol. 36 No. 3, pp. 1092-1113. https://doi.org/10.1016/j.ijforecast.2019.11.005

Peykani, P., Sargolzaei, M., Tanasescu, C., Shojaie, S.E. and Kamyabfar, H. (2025), 'Investigating the relationship between liquidity risk, credit risk, and solvency risk in banks listed on the Iranian capital market: A Panel Vector Error Correction Model', Economies, Vol. 13 No. 5. https://doi.org/10.3390/economies13050139

Safiullah, M. (2021), 'Financial stability efficiency of Islamic and conventional banks', Pacific Basin Finance Journal, Vol. 68. https://doi.org/10.1016/j.pacfin.2021.101587

Sakouvogui, K. (2020), 'Impact of liquidity and solvency risk factors on variations in efficiency of US banks', Managerial Finance, Vol. 46 No. 7, pp. 883-895. https://doi.org/10.1108/MF-05-2019-0241

SBP (2025a), Financial Stability Review 2024, State Bank of Pakistan, Karachi.

Sendi, A., Banna, H., Hassan, M.K. and Huq, T.I. (2024), 'The effect of ESG scores on bank stability: Islamic vs. conventional banks', Journal of Sustainable Finance and Investment. https://doi.org/10.1080/20430795.2024.2395876

Sharma, P., Mishra, B.B. and Rohatgi, S.K. (2023), 'Revisiting the impact of NPAs on profitability, liquidity and solvency: Indian banking system', IMIB Journal of Innovation and Management. https://doi.org/10.1177/ijim.221148863

Sugozu, I.H., Verberi, C. and Yasar, S. (2025), 'Machine learning approaches to credit risk: Evaluating Turkish participation and conventional banks', Borsa Istanbul Review, Vol. 25 No. 3, pp. 497-512. https://doi.org/10.1016/j.bir.2025.02.001

Tavana, M., Abtahi, A.R., Di Caprio, D. and Poortarigh, M. (2018), 'An artificial neural network and Bayesian network model for liquidity risk assessment in banking', Neurocomputing, Vol. 275, pp. 2525-2554. https://doi.org/10.1016/j.neucom.2017.11.034

Uula, M.M. and Kassim, S. (2024), 'Machine learning in Islamic finance: A bibliometric review', Dimensions-based literature review.

Vuong, G.T.H., Phan, P.T.T., Nguyen, C.X., Nguyen, D.M. and Duong, K.D. (2023), 'Liquidity creation and bank risk-taking: Evidence from a transition market', Heliyon, Vol. 9 No. 9. https://doi.org/10.1016/j.heliyon.2023.e19141

Zhu, Y., Xie, C., Wang, G.J. and Yan, X.G. (2016), 'Predicting China's SME credit risk in supply chain finance based on machine learning methods', Entropy, Vol. 18 No. 5. https://doi.org/10.3390/e18050195

Akram, H. and Hushmat, A. (2026), 'Bank liquidity creation, loan concentration and liquidity risk: a comparative analysis of dual banking system', Journal of Islamic Accounting and Business Research, Vol. 17 No. 1, pp. 179-201. https://doi.org/10.1108/JIABR-07-2023-0228

SBP (2025b), Islamic Banking Bulletin: June 2025, State Bank of Pakistan, Karachi.