A Hybrid Markov Decision Process (MDP) Model for Predicting Liquidity Profiles of US Regional Banks

by Khoa Nguyen

B.A. in Mathematics and Economics, May 2011, Colby College
M.B.A in Finance and General Management, May 2019, Columbia Business School

A Praxis submitted to

The Faculty of
The School of Engineering and Applied Science
of The George Washington University
in partial fulfillment of the requirements
for the degree of Doctor of Engineering

January 9, 2026

Praxis directed by

Amir Etemadi
Associate Professor of Engineering and Applied Science

A Hybrid Markov Decision Process (MDP) Model for Predicting Liquidity Profiles of US Regional Banks

Khoa Nguyen

Praxis Research Committee:

John Fossaceca, Adjunct Professor of Engineering and Applied Science, Praxis Chair

Amir Etemadi, Adjunct Professor of Engineering and Applied Science, Praxis Advisor

Walid Hassan, Professorial Lecturer in Engineering and Applied Science, Committee Member

© Copyright 2025 by Khoa Nguyen
All rights reserved

Dedication

To my father Nguyen Quan Luan and my mother Vo Thi Bich Phuong, my partner for life Tieu Thao Tran and to all my beloved family members whose unwavering love, support, and sacrifices have been my guiding light and foundation throughout my life and academic journey.

To Dr. Amir Etemadi, whose mentorship, guidance, and encouragement have been an enduring source of inspiration. Your leadership and belief in my potential have profoundly shaped this work and my pursuit of excellence.

To Professor Joseph Stiglitz, Bruce Greenwald, Michael Donihue, and Peter Howe, whose scholarship and teaching have enriched my understanding of the world and contributed meaningfully to my intellectual formation.

To Shelby Davis, the Davis Foundation, Gruppo Intesa Sanpaolo, and the United World College of the Adriatic in Trieste, Italy. Your vision and generosity have transformed countless lives and instilled in me a lifelong belief in the power of education to create positive change.

To my mentors Peter Babej, Lee Waite, Rob Goldstein, Patrick Marsh, and Shahmir Khaliq at Citigroup, Mark Howard at BNP Paribas, and Anne Clarke Wolff at Independence Point Advisors. Your wisdom, integrity, and dedication to excellence have been guiding examples throughout my professional life.

To my dear friends, whose kindness and presence have brought light, joy, and balance along this journey.

This thesis is dedicated to all of you, with profound gratitude for your faith, inspiration, and enduring support.

Acknowledgements

I would like to express my deepest gratitude to my advisor, Professor Amirhossein Etemadi, for his invaluable guidance and support throughout my doctoral research. His unwavering mentorship, thoughtful advice, and professional insight have been instrumental in the successful completion of this praxis. Without his constant encouragement and constructive feedback during each stage of this work, this achievement would not have been possible.

I extend my heartfelt thanks to my professors and the esteemed members of my thesis committee at The George Washington University, whose inspiring teaching, intellectual rigor, and insightful perspectives have profoundly shaped my academic journey and contributed significantly to the depth and quality of this study.

Finally, I wish to express my sincere appreciation to my colleagues at Citigroup for their steadfast support, encouragement, and camaraderie throughout this process. Their belief in my abilities and continued motivation has been a lasting source of strength and inspiration along this journey.

Abstract of Praxis

This praxis addresses the critical issue of accurately predicting liquidity for US regional banks during refinancing cycles. Regional banks are a vital part of the local economy, and they suffer more risk than larger financial institutions primarily because they serve a regional market, are less diversified structurally, and borrow from a limited number of potential sources. Inherent in this environment is increased risk exposure as they have limited access to capital should the local economy decline. Present liquidity prediction methodologies failing to consider the complexity of the nonlinear, spatial and temporal characteristics of liquidity profiles, leaving regional banks highly vulnerable to misjudged liquidity needs and refinancing costs, and financial instability.

This research proposes to reduce the predictive void by presenting a hybrid prediction framework based on Markov Decision Processes (MDP) and advanced machine learning algorithms in Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Long Short-Term Memory (LSTM). This framework leverages the structured decision-making capabilities of MDP and the robust predictive power of machine learning to enhance accuracy and timeliness in liquidity forecasts. The model is validated using liquidity profiles of regional banks listed in the S&P Regional Banks Select Industry Index (KRE) from 2013 to 2023, focusing on key macroeconomic variables such as interest rate spreads, credit spreads, and bond yield volatility. The hybrid framework couples machine learning accuracy with the MDP's dynamic decision optimization. Its advantage lies in policy performance, not raw predictive lift, allowing earlier intervention and improved liquidity resilience during refinancing cycles.

Table of Contents

Dedication iv

Acknowledgements v

Abstract of Praxis vi

List of Figures xiii

List of Tables xiv

List of Symbols xv

List of Acronyms xvii

Chapter 1—Introduction 1

1.1 Background 1

1.2 Research Motivation 2

1.3 Problem Statement 3

1.4 Thesis Statement 4

1.5 Research Objectives 4

1.6 Research Questions and Hypotheses 5

1.7 Scope of Research 6

1.8 Research Limitations 7

1.9 Organization of Praxis 7

Chapter 2—Literature Review 9

2.1 Introduction 9

2.2 Foundations of Liquidity Risk and Refinancing Cycles in Banking 10

2.2.3 Regulatory Responses and Structural Vulnerabilities 13

2.3 Traditional Approaches to Asset-Liability Management (ALM) 14
2.4 Rise of Machine Learning Models in Liquidity Prediction 17
2.4.1 Random Forest (RF) 18
2.4.2 Extreme Gradient Boosting (XGBoost) 19
2.4.3 Long Short-Term Memory Networks (LSTM) 20
2.4.4 Comparative Summary of ML Methods 20
2.5 Markov Decision Process (MDP) and Reinforcement Learning (RL) in Bank Liquidity Optimization 22
2.5.1 Theoretical Foundation and Intuition 22
2.5.2 Applications in Financial Decision-Making 24
2.5.3 Reinforcement Learning (RL): Learning from Interactions 24
2.6 Integrating MDP with ML: The Case for Hybrid Intelligence in Liquidity Management 26
2.6.1 Complementary Strengths of ML and MDP 27
2.6.2 Empirical Evidence Supporting Hybrid Approaches 28
2.6.3 Application of Hybrid Frameworks in Regional Banks 29
2.7 Literature Gaps and the Timeliness of this Praxis 30
2.7.1 Absence of Refinancing-Cycle-Awareness Frameworks 30
2.7.2 Underutilization of Macroeconomic and Sentiment Signals 31
2.7.3 Limited Model Interpretability for Supervisory Use 31
2.7.4 Inaccessibility for Mid-Sized Institutions 32
2.8 Summary and Transition to Methodology 32
Chapter 3—Methodology 34

3.1 Introduction to Methodological Framework 34
3.2 Data Sources and Integration 36
3.3 Feature Engineering and Label Construction 38
3.3.1 Market Microstructure Liquidity 38
3.3.1.1 Composite Liquidity Score 40
3.3.2 Balance Sheet and Capital Structure 40
3.3.3 Macro and Volatility Factors 42
3.3.4 Regimes and Event Time 43
3.3.5 MDP-Derived State Features 44
3.3.6 Temporal Derivatives and Interactions 46
3.3.7 Targets for Classification and Severity 47
3.3.8 Normalization, Leakage Control, and Drift 49
3.4 Model Classes and Training Protocol 50
3.5 Calibration, thresholds, and validation splits 52
3.6 Regularization, hyperparameters, and ablations 53
3.7 Diagnostics: Stability, Variance Inflation Factor (VIF), and Drift Readiness: 53
3.8 Explainability Assets (Importance, SHAP): 54
3.9 MDP State Space and Transition Estimations 54
3.10 Reward, Actions, and Constraints 56
3.11 Policy learning (DQN) and evaluation horizon H4 (four-quarter forward decision horizon) 57
3.12 Hypothesis testing strategy and statistical controls 62
3.13 Summary of Methodological Rigor 64

Chapter 4—Results 66
4.1 Introduction 66
4.1.1 Sample Declaration 66
4.1.2 Univariate Distributions 68
4.1.3 Correlation Structure 69
4.1.4 Summary Statistics (variables used downstream) 70
4.2 Predictive Model Performance 70
4.2.1 Evaluation Protocol and Metrics 71
4.2.2 Aggregate Test Metrics 71
4.2.3 Accompanying Figures 71
4.2.4 Construction Notes 72
4.3 Explainability Artifacts 72
4.3.1 Global Importance - tree ensemble 73
4.3.2 Global SHAP summary - tree ensemble with MDP inputs 74
4.3.3 Event-window contributions - YieldVol specification 75
4.4 Event-time outputs around refinancing peaks 75
4.4.1 Event-time setup 76
4.4.2 Figures 76
4.4.3 Event-time tables 77
4.5 Timeliness and Incremental Value 78
4.5.1 Definitions and Evaluation Setup 78
4.5.2 Supporting Figures 79
4.6.2 Transition Dynamics 80

4.6.1 Overall Transition Matrices 81
4.6.2 Yearly Transition Probabilities 83
4.7 Regime-Conditioned Transition Dynamics 85
4.7.1 Regime Definitions and Sample 86
4.7.2 Regime-conditioned persistence tables 86
4.7.3 Figures 87
4.8 Econometric Outputs: Fixed-Effects OLS and Diagnostics 88
4.8.1 Model Specification and Sample 88
4.8.2 Fixed-effects OLS results (binary outcome) 89
4.8.3 Fixed-effects OLS results (continuous outcome) 91
4.8.4 Multicollinearity diagnostics (VIF) 92
4.9 Scenario Outcomes and Resilience 93
4.9.1 Step-level trajectories (H4) 93
4.9.2 Scenario Totals 95
4.9.3 Bank-Level Resilience Panels 97
4.10 Model Monitoring: Drift and Retraining Outcomes 98
4.10.1 Population Stability Index (PSI) time series 98
4.10.2 Retraining Triggers 99
4.10.3 PSI panel by bank and quarter 101
4.11 DQN Policy Results 102
4.11.1 Reward-Weight Ablation (Episode Value and Trade-Off) 103
4.11.2 Policy effect on risk probability (H4 steps) 104
4.11.3 Step-Level Contributions (H4) 105

Chapter 5—Discussion and Conclusions 106

5.1 Discussion 106

5.2 Conclusions 110

5.3 Contributions to the Body of Knowledge 111

5.1 Recommendations for Future Research 112

References 114

Appendix A 121

Appendix A-1 121

Appendix A-2 121

List of Figures

Figure 4-31. DQN reward-grid trade-offs (means across banks) 103

Figure 4-33 H4: average P(low-liquidity) - base vs policy..... 104

Line chart of p_low_base and p_low_policy averaged across banks by forecast step $k \in \{1,2,3,4\}$. 104

Figure 4-34 H4 step contributions (mean across panel) 105

List of Tables

Table A-1. LSTM Diagnostics: Sequential Signal and Baselines 121

List of Symbols

$\hat{P}^{\text{hist}}$	Empirical Dirichlet-smoothed transition kernel (historical)
$P_{i,t}^{\text{high,hist}}$	Historical probability of entering high-risk state given previous state
$\hat{P}_{\alpha}$	Dirichlet-smoothed transition estimator (posterior mean)
$P_{i,t}^{\text{high}}$	One-step probability of high state from smoothed kernel
$y_{i,t}^{(k)}$	Lower-tail classification target at horizon k
$y_{i,t}^{\text{sev},(k)}$	Severity target (continuous shortfall below quantile threshold)
$y_{i,t}^{\text{LCR\_gap},(k)}$	Forward LCR shortfall target
$q_{\tau}(\cdot)$	Cross-sectional τ-quantile threshold
F_1	F1 score
$t^{\star}$	Threshold maximizing F1
Cost-aligned t^{*}	Threshold maximizing expected reward with costs (α, β, γ)
s_t	MDP state (s_t^{liq}, z_t)
z_t	Regime vector (e.g., RefiWindow, StressRegime)
A	Action set = {Hold, IssueTerm, SellHQLA, DrawRepo}
$R(\cdot)$	One-step reward function
CVaR	Conditional Value at Risk (tail-risk criterion)
$\sigma(\cdot)$	Logistic link function
ϕ_j	SHAP value for feature j
$z_r(\cdot)$	Robust z-score using MAD

$\Delta x_{i,t}$	Quarter-over-quarter change
$\mathrm{Slope}_L(x)$	L-quarter rolling slope
$[x]_+$	Positive part operator $\max\{x, 0\}$
$\mathbb{1}\{\cdot\}$	Indicator function
α	Dirichlet prior strength; also cost/weight parameters (context-dependent)
γ	Discount factor (e.g., 0.95)
λ	Regularization or penalty weight; also $\lambda_{\mathrm{LCR}}, \lambda_{\mathrm{NSFR}}$
θ	Risk-aversion parameter (entropic utility)
ϵ	Exploration rate; also small clipping constant
η	CVaR auxiliary variable
μ	Behavior policy (for OPE)
$\rho_{0:t}$	Per-decision importance-sampling weight
W_t	Index set of sequence window $\{t - L + 1, \ldots, t\}$
$X_{i,t-L+1:t}$	Feature window fed to sequence models
$p_{i,t+1}$	Predicted probability of next-state class
$\hat{p}, p^*$	Uncalibrated and calibrated probabilities
$\mathrm{Gini}(S)$	Gini impurity of set S

List of Acronyms

DHS	Department of Homeland Security
DoD	Department of Defense
U.S.	United States
MDP	Markov Decision Process
ML	Machine Learning
RL	Reinforcement Learning
RF	Random Forest
XGBoost	Extreme Gradient Boosting
LSTM	Long Short-Term Memory (neural network)
DQN	Deep Q-Network
ALCO	Asset–Liability Committee
SR 11-7	Federal Reserve Supervisory Guidance on Model Risk Management
OCC 2011-12	OCC Supervisory Guidance on Model Risk Management
TRACE	Trade Reporting and Compliance Engine
FFIEC	Federal Financial Institutions Examination Council Call Report forms
CBOE	Chicago Board Options Exchange
VIX	CBOE Volatility Index
PCA	Principal Component Analysis
PC1	First Principal Component
MAD	Median Absolute Deviation
HHI	Herfindahl–Hirschman Index (here: funding concentration)

HQLA	High-Quality Liquid Assets
LCR	Liquidity Coverage Ratio
NSFR	Net Stable Funding Ratio
CET1	Common Equity Tier 1 Ratio
RWA	Risk-Weighted Assets
LDR	Loan-to-Deposit Ratio
UST	U.S. Treasuries
CDs	Certificates of Deposit
PR AUC	Precision–Recall Area Under the Curve
AUC	Area Under the ROC Curve
VIF	Variance Inflation Factor
SHAP	SHapley Additive exPlanations
TreeSHAP	SHAP algorithm specialized for tree models
PSI	Population Stability Index
KS	Kolmogorov–Smirnov (two-sample test)
OPE	Off-Policy Evaluation
HAC	Heteroskedasticity- and Autocorrelation-Consistent (Newey–West)
FDR	False Discovery Rate (Benjamini–Hochberg)
DM test	Diebold–Mariano forecast comparison test
H4	Four-quarter decision/evaluation horizon
ATQ	Total Assets (quarterly)
SEQQ	Total Equity (quarterly)
DLTTQ	Long-Term Debt (quarterly)

DLCQ	Short-Term Debt (quarterly)
DPQ	Depreciation (quarterly)
LiqScore	Composite liquidity score (unit-weighted or PCA-weighted)
ILLIQ	Amihud Illiquidity measure
PriceStd	Standard deviation of execution prices (quarterly price dispersion)
Roll	Roll's effective spread estimator
HighLiquidity	Binary contemporaneous state (above cross-sectional median)
ST Funding	Short-Term Funding Ratio (DLCQ / Liabilities)
RefiLoad	Refinancing burden over next K quarters
CreditSpread	Median secondary-market spread (e.g., BBB vs A financials)
RefiWindow	Refinancing pressure indicator (volume/near-term maturity)
StressRegime	Market stress indicator (VIX or yield volatility > mean + 1σ)
quarter_key	Time index aligning datasets to bank-quarter
K	Default near-term horizon in quarters (often K = 4)
Event-time k	Alignment index around refinancing peak (k = 0 at peak)

Chapter 1—Introduction

1.1 Background

Regional banks in the United States play an important one as the custodians of local deposits and as the efficient suppliers of credit to small businesses, to households, and to municipalities. Being the link between small-scale financial service providers and the larger capital markets, these banks serve an especially important function in regional economies. At the same time that they serve valuable functions, however, the specialized nature of regional banks also exposes them to financial instabilities, particularly in refinancing periods, given their less diversified sources of funding as compared to larger financial institutions (Bartel, Hanke, & Petric, 2024).

The S&P Regional Banks Select Industry Index (KRE) serves as a useful benchmark to assess the performance of US regional banks, stressing the important role they play in providing financial stability to the broader banking network. Liquidity crises that have recently hit the regional banks sector highlight the necessity of more advanced tools to predict and manage liquidity risks. This praxis develops a hybrid forecasting-and-decision framework wherein machine-learning (ML) models (Random Forest (RF), Extreme Gradient Boosting (XGBoost), Long Short-Term Memory (LSTM)) forecast with reasonable accuracy near-term bank-level liquidity states and risk, while a Markov Decision Process (MDP) prescribes, within a set of certain regulatory-constrained conditions, state-contingent Treasury policy actions (e.g., term issuance, HQLA sales, repo draws) that directly trade off liquidity protection and costs. In essence, the probabilities are generated by ML; the probabilities are converted into actions via MDPs.

The integration model aims to accomplish two goals: (i) the timely and accurate prediction of liquidity stresses coming out of refinancing cycles, and (ii) the optimization of the policies which turn those forecasts into repeatable and auditable playbooks for Asset and Liability Committee (ALCO).

The synergy of MDP with these ML methodologies provides a powerful structure. MDP argues for the methods of optimizing decisions sequentially under uncertainty (Sutton and Barto, 2018), while XGBoost, RF and LSTM applied offers enhanced predictive power through the extract of complex nonlinear and time changing relations. The combination of these hybrid methodologies delivers a more robust prediction of liquidity dynamics.

1.2 Research Motivation

The Deposit Insurance Fund charged all surviving banks in 2024 $16.3 billion ($16.3Bn) as special assessments to make good all the losses incurred as a result of liquidity crises and the closure of banks during 2023 (The Federal Deposit Insurance Corporation, 2023). The liquidity risks of regional banks, if not properly regulated, can lead to systemic instability. Errors in the liquidity assessments raised the costs of refinancing but also depressed the solvency of regional banks, thereby impairing the investor and regulatory confidence placed in those institutions.

Conventional liquidity assessment models frequently fail to capture the funding dynamic and localized risks that regional banks face due to their inherent attitudes toward static assumptions and lack of adaptability and real-time response. Using the decision-making structure of MDP and the predictability of the development of machine learning techniques, this research seeks to provide a more robust and dynamic solution. MDP

provides a sequential decision model to enhance the efficiency of XGBoost, which is optimized for large and more complex datasets, and accompanies the ensemble strengths of RF in reducing overfitting and LSTM in capturing time series dependencies. With this hybrid architecture, this praxis aims to provide stakeholders with actionable intelligence that can help improve decision-making, reduce risk, and fortify the resilience of regional banks.

1.3 Problem Statement

Inaccurate liquidity assessment degrades regional banks' performance and can precipitate bank runs around refinancing cycles (Kashyap, Rajan, & Stein, 2002), imposing real costs such as the $16.3 billion special assessment to replenish the Deposit Insurance Fund in 2023 (Federal Deposit Insurance Corporation, 2023)

The problem addressed in this research is that U.S. regional banks face significant challenges in accurately predicting liquidity during refinancing cycles, which can result in severe financial instability. These banks are especially vulnerable because of their limited funds supply and orientation toward regional economies, increasing the likelihood that their liquidity will be miscalculated. When liquidity prediction failures occur, regional banks suffer a chain of events, including rising refinancing costs, regulatory scrutiny, and decreasing confidence from stakeholders involved. The existing tools currently in use do not deal easily with the complex, nonlinear, time-dependent terrain of liquidity, since they deal with static assumptions that cannot adapt to changing, if not fast-changing, market conditions. Because liquidity decisions are partly determined by market situations, banks that can predict their liquidity plans accurately during refinancing periods will be able to give themselves a competitive advantage in handling

their liquidity problems and increasing financial stability for themselves and the broader banking system.

1.4 Thesis Statement

A hybrid model in which a machine-learning forecaster estimates regional bank liquidity states around refinancing cycles, and a Markov Decision Process uses those estimates to optimize Treasury actions under costs and regulatory constraints.

The principal idea in this research is that a hybrid predictive framework based on MDP, and sophisticated modern ML models will generate considerable increases in the accuracy and timeliness of the liquidity predictions for regional banks. Regional banks face complex liquidity dilemmas produced by changing environmental variables and refinancing cycles, which will result from changing environmental conditions that can be difficult to manage with conventional techniques. Even the treasury manager and the experienced senior decision-makers will find substantial restrictions on their ability accurately to predict liquidity requirements, particularly in circumstances of changeable macroeconomic factors. A framework is suggested in the research which will automatically process the complex liquidity datasets which are characteristic of forecasting exercises which involve sequential decision-making and non-linear characteristics improving the predictions' accuracy. This framework will provide a very powerful and scalable liquidity management framework for the stakeholders to deal with the complex problems of liquidity management faced.

1.5 Research Objectives

This praxis introduces a hybrid framework that couples ML forecasts with MDP. In preview of later results, the inclusion of MDP state features does not materially alter

out-of-sample discrimination in the present data set, but the integrated framework provides probabilities usable at the implementation level and supports scenario rollups, monitoring, and decision discipline throughout refinancing cycles. To achieve this goal, the research has been divided into the following sub-objectives: (1) to determine the more salient factors influencing the liquidity dynamics of the subject population through the use of feature selection alternatives such as generating features, feature importance implementation, correlation heat map; (2) to compare and evaluate the effectiveness of various ML methods: RF, XGBoost and LSTM, in combination with MDP and thus be able to designate the best hybrid architecture; (3) to generate and test a predictive system coupling MDP with the best ML model to provide to regional banks with timely and accurate liquidity decisions required during refinancing cycles.

1.6 Research Questions and Hypotheses

This praxis explores the effectiveness of a hybrid model that combines MDP with ML techniques in predicting liquidity for regional banks. The research questions and hypotheses are below:

RQ1: Among RF, XGBoost and LSTM, XGBoost when combined with MDP would most accurately predict liquidity for regional banks during refinancing cycles?

H1: With the identical "with MDP" feature set, XGBoost (with MDP) achieves higher out-of-sample AUC (and F1) than Random Forest (with MDP) and LSTM (with MDP) during refinancing windows.

RQ2: How do rising interest rates and widening credit spreads impact regional banks' liquidity levels during their refinancing cycles?

H2: Higher rate levels, higher rate volatility, and wider credit spreads are associated with lower liquidity for regional banks, with effects amplified inside refinancing windows.

RQ3: Does the learned policy π minimize time-in-risk compared to a baseline non-optimized liquidity strategy?

H3: The learned policy function π reduces cumulative time-in-risk relative to the baseline, demonstrating the value of sequential decision optimization

1.7 Scope of Research

The KRE index serves as a solid benchmark for representatives of the regional banks of the US. Because of the focus on liquidity, this praxis will closely examine liquidity profiles during the refinancing cycles from 2013 to 2023 for each bank represented in the index. This praxis works from pricing datasets from bonds, preferred securities, and convertible securities, issuance schedules which are obtainable on the bank level, together with volatility trends, and interest rates as part of the macro environment.

The two data sets serve as base inputs into the hybrid model in this praxis. As part of praxis implementation, the MDP is given two complementary functions. The first estimates empirical state transitions $P(S_{t+1} \mid S_t)$using panel data, and the resulting transition features are used to improve supervised prediction (as detailed in Chapter 4, §§4.1 - 4.6). The second trains a Deep Q-Network (DQN) policy in the same state space to evaluate decision sequences given defined actions and reward structures (§3.5). (The two streams are subsequently coupled through the empirical transition kernel introduced

in §3.11, ensuring that the policy layer operates on the same data-driven dynamics inferred by the predictive layer.)

1.8 Research Limitations

Regional banks encounter an unusual challenge due to their funding model. The limitations below are purposely placed to ensure we maintain focus for this analysis:

1.The hybrid framework has been designed for U.S. regional banks in the KRE index, it has not been tested against the larger national banks, credit unions, or non-U.S. financial institutions limiting its generalizability.

2.The research is focused on liquidity predictions during refinancing cycles through a defined period (2013-2023). As noted, financial crises tend to develop over time, and the model may not be entirely successful in all financial environments through varying cycles of interest rate conditions perceived.

3.The “black box" nature of machine learning (ML) models could limit the transparency of the actual decision-making process.

1.9 Organization of Praxis

This praxis creates a combined framework of MDP and ML to predict and enhance liquidity management in regional banks across the United States during refinancing periods. The model incorporates supervised learning techniques such as RF, XGBoost, and LSTM, alongside an MDP optimization component that assesses sequential decisions made under conditions of liquidity pressure.

Chapter 2 reviews literature on liquidity risk forecasting, asset-liability management (ALM), ML and reinforcement learning (RL) in banking. Existing approaches to predicting liquidity are reviewed, including works on MDP, RF, XGBoost,

and LSTM models. Limitations of standalone approaches and benefits of integrating MDP with ML techniques will be examined.

Chapter 3 outlined the research methodology thoroughly. A step-by-step approach is laid out to convey the work that was performed regarding data analysis, including data sources, preprocessing, feature extraction, model building and validation.

Chapter 4 displays the results of the study regarding liquidity analysis in the form of tables and graphs. Also contained in this chapter are all the significant findings of the study which cover the efficiency of the hybrid model to predict liquidity profiles of regional banks in comparison with traditional models and the other individual ML techniques.

Chapter 5 describes the Discussion and Conclusions of the thesis. This chapter is devoted to showing the practical applications of the hybrid MDP and ML model to liquidity forecasting and liquidity optimization and its importance to actual decisions. The limitations of the study are explained with recommendations on how the study can develop in future work. The chapter ends with a summary of the contribution of the work presented to treasury and liquidity management, with a focus on effectiveness, interpretability and scalability of the model to the management of liquidity risks through funding and refinancing cycles.

Chapter 2—Literature Review

2.1 Introduction

Interest rates play an important role in bank's funding and balance sheet management. Given a significant fixed income securities portfolio in each regional bank's balance sheet, banks can rapidly become insolvent by virtue of holding on to long dated assets in a rising rate environment.

2023 was a continuation of an upward trend in rates, following a series of rate hikes by the Fed that started in mid-2022. With this rapid rise in rates, many of the regional banks were caught off guard. Notable among them were Silicon Valley Bank, Signature Bank, and First Republic Bank. Traditional liquidity risk models that were static did not provide the timely insights needed under these dynamic stress conditions. Recent advances in sequential decision-making models, predictive analytics, and Reinforcement Learning (RL) have become building blocks to help establish a more holistic approach to modern liquidity management. In subsequent sections, this praxis will walk through relevant theoretical frameworks to contribute to the overall model.

Section 2.2 examines the theoretical underpinnings and regulatory context of liquidity risk and refinancing cycles. Section 2.3 analyzes conventional asset-liability management (ALM) methods. In Section 2.4, the focus shifts to the rise of machine learning (ML) in liquidity forecasting, along with its associated constraints. Section 2.5 introduces the Markov Decision Process (MDP) and RL as tools for dynamic decision-making. Section 2.6 delves into emerging hybrid frameworks that integrate MDP with machine learning (ML) techniques. Section 2.7 consolidates the key gaps observed across

the literature. Section 2.8 offers a concluding synthesis while laying the groundwork for the methodology chapter.

2.2 Foundations of Liquidity Risk and Refinancing Cycles in Banking

2.2.1 The Nature of Liquidity Risk

Banks serve inherently as duration transformers in that they collect short-term deposits to lend long term. Duration mismatch provides a challenge to either type of bank, global banks headquartered in New York and local regional banks throughout the country. However, while Global Systemic Important Banks (GSIB's) have been structurally reinforced via diversified funding sources, ability to access central bank liquidity and sophisticated liquidity contingency plans, particularly during times of stress or financial turmoil, regional banks are subjected to dramatically different restraints, their liquidity buffers being inherently less diverse, often being condensed into less liquid asset categories such as commercial real estate and municipal securities. Additionally, their dependence upon wholesale funding markets, brokered deposits, and lack of institutional diversification, renders them particularly vulnerable to funding pressures and abrupt shifts in market sentiments. One commonly accepted definition of liquidity risk occurs in the context of the Bank of International Settlement (Bank for International Settlements [BIS], 2008) is the fundamental liquidity mismatch existing between the time phase, quantity and certainty of a financial entity's incoming cash flows vs. its short-term liabilities.

One common definition for liquidity risk, according to the Bank of International Settlement (Bank for International Settlements [BIS], 2008), is the fundamental mismatch between the timing, magnitude, and certainty of a financial institution’s cash inflows and

its short-term obligations. Liquidity risk is generally classified into two interdependent dimensions: 1) funding liquidity risk, in essence the inability to renew maturing liabilities or procure new funding under stress, and 2) market liquidity risk, denoting the difficulty of liquidating assets rapidly without incurring material value losses (Drehmann & Nikolaou, 2013).

These two liquidity dimensions are obviously interdependent and not merely additive but indeed are non-linear in orientation during periods of financial stress, thus exacerbating systemic fragility. Funding deficiencies can force asset liquidation, which can drive values lower in an illiquid environment, thus reducing capital, increasing funding pressures, and further perpetuating funding deficiencies, which can create a self-reinforcing and dynamic feedback loop often underestimated in traditional models (Acharya & Mora, 2015).

2.2.2 Refinancing Cycles as Points of Stress

Refinancing cycles represent concentrated time windows during which a substantial proportion of a bank's longer-term liabilities, such as senior bonds, structured deposits, certificates of deposit, and commercial paper, reach maturity and require rollover. These episodes inherently elevate a bank's exposure to timing mismatches, funding cost volatility, and systemic liquidity risk (Gorton & Metrick, 2012).

Refinancing cycles therefore can easily introduce periodic stress points into the liquidity structure. If market conditions deteriorate concurrently to a refinancing event, banks may be forced to accept less favorable funding terms, liquidate assets at fire-sale prices, or resort to emergency liquidity facilities at the Fed and other banks, each action of which will impose escalating costs and regulatory scrutiny.

Empirical evidence from 2023 shows how an acute deposit outflow can collapse the effective maturity of bank liabilities and thereby *pull forward* a refinancing window. Let L_t^{dep} denote runnable deposits and B_t^{wh} wholesale debt with contractual maturities $\{T_j\}$. A rapid deposit run $\Delta L_t^{\mathrm{dep}} < 0$ creates an immediate funding gap that must be closed through (i) asset liquidation, (ii) secured funding at higher haircuts/spreads, or (iii) term issuance earlier than planned. Simultaneously, mark-to-market losses on long-duration securities reduce HQLA and collateral value, raising both new-issue concessions and repo haircuts. The result is that a deposit shock compresses the liability ladder, converts near-term uncontracted outflows into an unscheduled refinancing requirement, and amplifies liquidity risk through the channels that this praxis aims to model and optimize.

A simple accounting identity makes the flow above explicit:

$$\underbrace{FundingGap_{t,h}}_{\text{Net gap to be financed}} = \underbrace{\Delta L_{t,h}^{dep}}_{\text{gross outflow to be financed}} - \underbrace{\sum_{j=1}^{J} B_{t,j}}_{\text{scheduled roll}} - \underbrace{SaleableHQLA_t}_{\text{post-haircut capacity}} + \underbrace{SecuredCapacity_t}_{\text{post-haircut inflow}}$$

When $\mathrm{FundingGap}_t > 0$, the bank must issue term funding within horizon h, thereby entering a refinancing window earlier than implied by contracts maturities. In 2023, institutions with large unrealized losses and concentrated uninsured deposits exhibited precisely this mechanism: bank runs shrank effective maturity, collateral values fell, market spreads widened, and refinancing (or liquidation) had to occur under stress pricing (Jiang et al., 2023; Drechsler et al., 2021).

Process schematic:

1. Deposit shock $\downarrow L_t^{\mathrm{dep}} \rightarrow$ immediate outflow (Diamond & Dybvig, 1983; Iyer & Puri, 2012).
2. Market value shock ↓HQLA/collateral → higher haircuts, lower secured capacity (Adrian & Shin, 2010; Gorton & Metrick, 2012).

3. Gap opens: $\text{FundingGap}_t > 0$.
4. Forced actions: Sell HQLA / Draw repo / IssueTerm (unscheduled) consistent with LCR/NSFR constraints (BCBS, 2013).
5. Refinancing window is pulled forward and amplifies liquidity risk via worse spreads and concessions (He & Xiong, 2012; Greenwood, Hanson, & Stein, 2015).

2.2.3 Regulatory Responses and Structural Vulnerabilities

In the wake of the Global Financial Crisis of 2007- 2009, the Basel Committee on Banking Supervision, through Basel III proposal, called for the Liquidity Coverage Ratio (LCR) and the Net Stable Funding Ratio (NSFR) to enhance banks' ability to withstand liquidity shocks (BCBS, 2010)[1]. These provisions integrated quantitative liquidity buffers and maturity funding stability into the supervisory framework. However, for many U.S. regionals, the effective range of those requirements remained limited through either partial exemptions or tailoring, which reduced bindingness relative to the largest institutions (Cornett, McNutt, Strahan, & Tehranian, 2011). As a result, key vulnerabilities persisted.

Many regional banks still depend upon legacy ALM structures as operationally defined by static assumptions, limited integration of real-time data and infrequent manual stress-testing processes. This friction inhibits timely measurement of funding risk and transmission of market shocks.

[1] The Liquidity Coverage Ratio (LCR) requires a bank to hold enough high-quality liquid assets to cover its net cash outflows over a 30-day stress period, ensuring short-term survival. The Net Stable Funding Ratio (NSFR) requires banks to maintain stable funding for their assets and commitments over a one-year horizon, promoting long-term resilience by incentivizing stable funding sources like equity and long-term debt over short-term wholesale funding. The LCR addresses short-term liquidity risk, while the NSFR addresses long-term structural funding risk. www.bis.org/publ/bcbs238.pdf

Evidence supports the materiality of these structural defects. Complementary analysis by Acharya, Bergant, Crosignani and McCann (2024) indicates that banks with high levels of uninsured reliance on deposits and concentrations of callable debt were subject to more rapid liquidity shortfalls during a stress period than comparable peer institutions. The refinance cycles revealed by these studies suggest not only localized roll over events, but amplification processes which interact with funding mix, depositor behavior, and market segmentation thereby enhancing liquidity depletion processes.

Altogether, the evidence shows that static compliance to LCR/NSFR, although necessary, is not sufficient. A next-generation approach would require liquidity analytics that are forward looking and include design of dynamic scenarios and prompt operational playbooks. More specifically, firms would benefit from models that (i) nowcast liquidity in market microstructure terms, (ii) forecast state transitions when refinancing opportunities arise and that (iii) convert risk evaluations to pre-committed actions in Treasuries that are consistent with supervisory expectations for model risk management and explainability (Lundberg & Lee, 2017).

2.3 Traditional Approaches to Asset-Liability Management (ALM)

Historically, when the traditional ALM frameworks were designed, the process was primarily to align the maturities, cash flows and interest rate sensitivities of a bank's assets and liabilities to lessen the bank's exposure to market valuation risks. Foundational concepts such as duration matching, gap analysis, and immunization theory were designed to develop a form of static equilibrium of the balance sheet (Redington, 1952; Bierwag, 1987). These concepts became so widely accepted due to the assumptions made

about stability in funding markets, stability in depositor behavior, and stability in the valuation of underlying assets.

The objective of duration matching is to equalize the weighted-average maturities (durations) of the assets and liabilities held by the institution to immunize the economic value of the bank against small parallel shifts in the yield curve. A gap analysis measures the mismatches between cash flows in specified time buckets. To hedge against arbitrary nonlinear movements in interest rates which are nonparallel, modern ALM has incorporated the concepts of convexity and key-rate duration for preserving the Economic Value of Equity (EVE) sensitivity to shocks and yield-curve twists (Bierwag, 1987; Tuckman and Serrat, 2011; BCBS, 2016/2021; Lubinska, 2021).

While these techniques could prove effective in environments where the liquidity conditions are stable, they sensibly have limitations in modern liquidity stress scenarios. First, they rest on the assumption of static balance sheet structure and ignore customer behaviors, embedded options (such as loan prepayment or deposit withdrawal options), and liquidity premium shifts. Second, they assume that movements in the yield curve will parallel those in the interest rate environment, ignoring the more complex yield curve twists and kinks which occur in these types of market disruption. Third, they abstract away the path dependence of liquidity demand and treat it much in the same way shocks to central bank liquidity are treated. In other words, liquidity experiences such as stress should not be treated as independent shocks and instead it would seem reasonable to assume that liquidity strain has the capacity to build and compound on successive episodes of market events (Borio, Drehmann, & Tsatsaronis, 2014).

In practice, traditional static ALM methodologies proved inadequate in anticipating liquidity crises during periods of financial stress. The 2023 regional bank failures exemplified this limitation, as concentrated refinancing exposures, rapid deposit outflows, and abrupt collateral revaluation shocks interacted in nonlinear and mutually reinforcing fashion, depleting liquidity buffers within days. Furthermore, these static frameworks lack the capacity for adaptive strategic adjustments, an essential requirement for managing liquidity in a stochastic and adversarial market environment characterized by high uncertainty and rapid regime shifts.

Following the Global Finance Crisis (GFC) 2008, regulatory reforms introduced scenario-based stress testing requirements to address some of these shortcomings (Dudley, 2017). Nevertheless, liquidity stress testing remains largely static and stylized: tests are performed infrequently, scenarios are limited to historical analogues, and institutions rarely update assumptions dynamically as market conditions evolve (Duffie & Krishnamurthy, 2016).

Infrastructure constraints amplify these issues. Many regional banks still operate liquidity risk management using stationary cash flow models based on spreadsheet design without having stochastic simulation engines, real time feeds of information or predictive analytics available (FDIC, 2023). The lack of real-time liquidity sensing and optimization of the balance sheet on a dynamic basis leaves the banks in a state of great vulnerability to unanticipated liquidity shocks.

In essence, while traditional ALM methodologies offer foundational insights into balance sheet risk structure, they are fundamentally insufficient for managing modern liquidity risks characterized by dynamic refinancing cycles, behavioral nonlinearities, and

systemic contagion channels. Accordingly, as liquidity shocks of a funding-based paradigm become more frequent and severe, the imperatives for liquidity management of an adaptive ML-enhanced decision theoretic nature become paramount.

2.4 Rise of Machine Learning Models in Liquidity Prediction

With the structural shortcomings of conventional ALM frameworks coming more clearly to light in recent high-volatility bouts and liquidity-induced refinancing stresses, financial institutions have begun to embrace ML techniques to enhance liquidity forecasting capabilities. ML methods, and in particular ensemble methods (e.g., RF and Gradient Boosted methods) and deep learning architectures (e.g., Long Short-Term Memory networks), are powerful in being able to model nonlinear, high-dimensional relationships prevalent in heterogeneous and large datasets, and often without the need for stringent parametric assumptions (Kelly & Xiu, 2023).

Such flexibility gives ML models the ability to discover latent patterns across both structured financial indicators (e.g., deposit growth, wholesale funding spreads, yield curve shifts), as well as unstructured sources of information, such as earnings call sentiment, media linkage sentiment, and macroeconomic stories. For regional banks with complex fundings that are less able to observe current market signals, these models mark a significant enhancement of predictive analytics (Sirignano & Cont, 2021).

Nevertheless, a core limitation still exists. Most of the ML models are predictive and not prescriptive, that is, they are optimally developed to anticipate negative liquidity events (e.g., increased withdrawal risk or stress spreads), without giving any insight into what actions a treasury desk should take strategically. This distinction is nontrivial. In complex financial systems, forecasting alone does not close the control loop. Without an

integrated decision-making capability, institutions may correctly anticipate liquidity stress yet still fail to act in time, leaving the loop of prediction, action, and feedback incomplete.

In contexts involving liquidity, ML models have shown great efficacy in predicting key risk variables including:

- Deposit outflows and flight risk exposures for banks with a high concentration of uninsured deposits or unrealized losses in securities as evidenced by the 2023 US banking turmoil (Jiang, Matvos, Piskorski, & Seru, 2023).
- Escalation in funding costs due to stress given increased credit spreads and continued volatility in wholesale funding in situations of liquidity crunches (Dudley, 2013).

Among the more popular implementations of techniques regularly used for these forecasting attributes are RF, XGBoost, and LSTM networks. Each of these have computational advantages and disadvantages with respect to liquidity risk applications, which are summarized below.

2.4.1 Random Forest (RF)

Random Forests (RF), which were introduced by Breiman (2001), are examples of ensemble learning models that combine the predictions of several decision trees that are built on randomized subsamples of data. Given its architecture, it has good resistance to overfitting phenomena and is especially useful in the presence of noisy or incomplete data, being known, for example, the case of bank liquidity reporting (Lessmann et al., 2015; Chakraborty & Joseph, 2017).

RF models are useful in assessing liquidity risk applications as they offer early warning signals through interaction effects across variables such as the composition of a firm's deposit base, market volatility, and interest rate movements. However, RF models remain static in nature, as they provide probabilities of liquidity stress at any point in time, but do not lend themselves to the modeling of intertemporal dependencies or sequence-sensitive liquidity dynamics. This limits their strategic utility in refinancing cycles where the evolution of distress is path-dependent and compounding.

2.4.2 Extreme Gradient Boosting (XGBoost)

Extreme Gradient Boosting (XGBoost), introduced by Chen and Guestrin (2016), builds on boosting by iteratively correcting the residual errors in earlier trees, while adding explicit regularization to control model complexity. This method has achieved wide acceptance in the financial domain, including default prediction, liquidity stress classification, and early warning diagnostics (Malekiipourzari & Aksakalli, 2015). It is fully explainable in this analysis by means of TreeSHAP. The SHAP framework provides additively consistent, local feature attributions for each prediction, and applies to all ensembles of decision trees, including both gradient boosting and bagging (Lundberg & Lee, 2017; Lundberg, Erion & Lee, 2020).

We compute SHAP summaries therefore for both the XGBoost and for Random Forest using identical explainer and feature-matrix. The practical recommendation of using XGBoost in our analysis derives from the method's boosting optimization, reducing bias on tabular data with mixed scales, not from any particular claim to explainability.

XGBoost remains a forecasting model rather than a sequential decision model. It has no built-in state dynamics and, by itself, does not capture the temporal evolution of risk. In the hybrid architecture proposed here, temporal dependence is captured through explicit state features and a Markov Decision Process (MDP) layer that consumes the model's probabilities during refinancing episodes, when the timing of intervention is critical.

2.4.3 Long Short-Term Memory Networks (LSTM)

Long Short-Term Memory networks (LSTMs) (Hochreiter & Schmidhuber, 1997) are explicitly designed for modeling long-range dependencies in sequences over time. They employ gated architecture that allows memory to be selectively retained, making them optimally suited for tasks in regard to liquidity prediction, which, by the nature of liquidity driven hypotheses, require temporal dependencies (for example, rolling deposit behavior, sequential bond maturity behavior, or volatility transmission from capital markets to bank funding costs).

In practice, LSTM models exhibit excellent performance in prediction of temporally dependent risks in such risk areas as asset volatility prediction or liquidity migration modeling (Fischer et al., 2018). But LSTM models do exhibit two significant drawbacks: they are very data-intensive, requiring large amounts of well labeled time-series data for training purposes, data which are often not going to be available at regional institutions, and they suffer from esotericism in that they are opaque relative to the procedures or algorithms they employ, making them challenging for difficult for challenges for regulatory auditability and supervisory model governance (Waqas & Humphries, 2024).

2.4.4 Comparative Summary of ML Methods

Summarized strengths and limitations of RF vs. XGBoost vs. LSTM in the modelling of liquidity risk:

Model	**Learning Mechanism**	**Regulatory Interpretability**	**Role in Hybrid MDP Pipeline**
RF	Ensemble bagging of decision trees	High - transparent, auditable	Baseline risk scorer and control benchmark
XGBoost	Sequential gradient boosting with regularization	Moderate - High with SHAP explainers	Core risk engine guiding MDP policy actions
LSTM	Recurrent network with gated memory for temporal learning	Low - Moderate; requires post-hoc explainers	Temporal signal generator for state transitions

Although these models have materially enhanced the financial sector's ability to predict emerging liquidity risks, they remain passive tools. They do not permit either optimization in real-time, or choice of policy. As a result, institutions equipped with early warning systems based on the ML model types used in the March 2023 banking turmoil, often remained unable to act assertively, despite useful predictive information. It is at this intersection between prediction and action, between knowing when a stress condition will occur, and knowing how best to respond to it, that MDP and RL offer the greatest transformational potential. MDP and RL methods allow the enhancement of decision

making over time, to reflect not only conditions as they evolve, but long-term trade-offs, and these are considered in more detail in Section 2.5.

2.5 Markov Decision Process (MDP) and Reinforcement Learning (RL) in Bank Liquidity Optimization

2.5.1 Theoretical Foundation and Intuition

An MDP is characterized by a state space S, an action space A, transition probabilities $P(s' \mid s, a)$, and a reward function $R(s, a)$ which indicates the immediate payoff that can be obtained from the action a when in state s (Puterman, 1994). In the area of liquidity management, the set of states may represent the liquidity position, the composition of funding or the state of the market for a bank; the set of actions may be related to refinancing, the buildup of reserves or the liquidation of assets or liabilities; while the reward function can be structured to penalize bad results, resulting from fire sales, excessive emergency borrowing, and out of regulatory limits. By considering liquidity management in MDP form, banks are able to dynamically optimize their funding and investment strategies under uncertainty, developing policies that result in a balance between short term liquidity maintenance and long-term profits under a regime of changing market conditions (Moody & Saffell, 2001).

In Treasury contexts, MDP measures the systematic effect of policy choices on future liquidity states and costs through systemic temporal dependence embodied in dynamic constraints. Recent applied work has extended the traditional MDP model class to richer settings that incorporate friction such as liquidity limits, issuance costs, and penalties, while maintaining a model framework with fully specified dynamics. When the transition equation or reward vector is only partially known or too complex to specify analytically,

the MDP to be solved requires approximations or data-driven substitutes for these quantities before standard dynamic programming methods can be applied.

Reinforcement learning (RL) relaxes the need for fully known dynamics in that it derives policies from data or simulations of those policies either by estimating the transition model (as in model-based RL) or directly optimizing of the value functions (as in model-free RL) (Sutton & Barto, 2018). In finance, more recent work has begun to study MDP-type problems by applying reinforcement learning (RL) to the optimal hedging under realistic market frictions. Huang and Lawryshyn (2025), for instance, model derivative hedging as a sequential-decision-making process by which the RL agent learns an optimal policy with respect to a continuous action space so as to minimize the hedging losses over time from transaction costs, slippage and permanent market impact. Their formulation operationalizes the structure of an MDP by defining the states as properties of the evolving market and portfolio values, actions as adjusted hedge decisions, and rewards as adjusted performance of the portfolio so as to demonstrate that RL can internalize liquidity and impact tradeoffs which are not available in closed form solutions.

In the case of liquidity management, reward functions in MDP frameworks can be structured to yield added penalties for bad results such as costly fire sales of assets, excessive emergency borrowing, or violations of regulatory liquidity coverage. As an illustration, liquidating securities during adverse market liquidity events could be given a very high negative reward in comparison to the rewards of preemptive refinancing or opportunistic liquidity reserve accumulation. By constructing rewards to expose real economic and regulatory results, MDP-based models can assist treasury groups in

ordering the urgency of proactive liquidity management policies in order to minimize financial, reputational, and supervisory costs over both short-term and long-term horizons.

2.5.2 Applications in Financial Decision-Making

MDPs have been effectively used in a variety of financial applications which show their generality in the sequential decision problems under uncertainty. Important examples include

- Increasing the stability and systemic risk reduction of interbank lending systems, by implementing optimally modelled intervention policies, for instance in relation to the bail-out policy (Capponi & Chen, 2015)
- Assistance in ALM problems, especially where trading or funding is constrained (Englisch et al., 2023)

MDP models rely heavily on the assumptions of stationary dynamics (the Markov property) and complete observability of the state. These restrictions are always difficult to satisfy in the current financial situation with their uniform non-stationarity, instability of market regimes, variability of volatility and lack of information. These inherent difficulties of MDP modelling involve a heavy model risk and have led to an accelerated use of DRL and other non-model related learning techniques. DRL offers greater flexibility by allowing models to learn optimal, adaptive strategies directly from interaction with complex, real-world environments.

2.5.3 Reinforcement Learning (RL): Learning from Interactions

RL enhances the MDP framework by allowing agents to learn policies via simulated interaction rather than requiring full knowledge of all characteristics of the environment

dynamics (Sutton & Barto, 2018). When combined with deep neural networks, this approach (known as Deep Reinforcement Learning) can be applied to high dimensionality problems, even in situations where the full state of the environment is not directly observable (Deng et al., 2016). Examples of Deep Reinforcement Learning applications in more complex financial environments include:

- Improving asset-liability management strategies against increasing interest rates (Wekwete et al., 2023).
- Constructing a model-free deep reinforcement learning framework using twin-delayed deep deterministic policy gradient algorithm in a mean-variance optimization environment, that achieves better dynamic portfolio allocation performance (Jiang, Olmo, & Atwi, 2024).

Despite its enormous potential, uptake of deep reinforcement learning has just recently begun within the banking environment. There remain important obstacles to be overcome such as the lack of transparency associated with its use, the limited quantity of data necessary to train the learning agent adequately, and the difficulty in ensuring that the policies developed satisfy regulatory requirements. An approach currently evolving is the use of predictive machine learning approaches linked with the general MDP or RL frameworks. That is a hybrid approach, involving predictions and structured decisions.

This praxis adopts a model-based formulation: a supervised learner estimates the one-step transition model, rather than overwriting the current state, $\hat{p}_\theta(s' \mid s, a)$ (or an action-free $\hat{p}_\theta(s' \mid s)$). The MDP layer then plans under these dynamics to obtain the value function and policy $V^\star(s) = \max_{a \in \mathcal{A}} \{R(s,a) + \gamma\, \mathbb{E}_{s' \sim \hat{p}_\theta(\cdot|s,a)}[V^\star(s')]\}, \pi^\star(s) = \arg\max_{a \in \mathcal{A}} \{\cdot\}$.

In short, ML supplies the transition dynamics; the MDP/RL solver supplies the

sequential action policy (Sutton & Barto, 2018). This separation keeps estimation and control modular, auditable, and aligned with supervisory expectations.

When actions influence next-state risk, this praxis uses an action-aware mixture kernel $P_{\alpha}(s' \mid s, a) = \alpha\, \hat{P}_{\text{emp}}(s' \mid s) \; + \; (1-\alpha)$ clip $(\tilde{p}_{t+1} + \Delta_a(s), \epsilon, 1-\epsilon)$, where $\tilde{p}_{t+1}$is the ML forecaster's probability and $\Delta_a(s)$is a first-order log-odds uplift mapped back to probability space, a standard linear approximation used in applied policy modeling. Δ_a here represents the model's assumption about the immediate marginal impact of an action, such as selling HQLA or issuing term debt, on the likelihood of transitioning to a stress state in the next period.

2.6 Integrating MDP with ML: The Case for Hybrid Intelligence in Liquidity Management

While ML models such as RF, XGBoost, and LSTM have made substantial improvements in liquidity risk detection, they nevertheless remain static in nature, providing insights into short term imminent stress events, but without regard for the optimization of sequential decision-making over time. MDP explicitly formalizes multi-stage decision problems under uncertainty, but requires fully specified transition dynamics, and reward functions, which are difficult if not impossible to fully specify in some classes of stochastic financial environments (Puterman, 1994).

The hybridization of ML models with multi-stage decision-making frameworks thus serves to unite the benefits of both archetypes; the ML models identify the onset of the early threat of a liquidity stress situation, while the MDP paradigm formalizes the optimization of strategic responses to them. This hybrid framework can thus offer a

unique and constructive approach to the settlements of dynamic adaptive liquidity management on the part of the regional banks.

2.6.1 Complementary Strengths of ML and MDP

The integration of ML and MDP addresses the core limitations of each technique when applied independently. Operationalizing such a hybrid framework within a regional bank treasury could proceed through the following architecture:

- Step 1: LSTM, RF or XGBoost models forecast near-term liquidity indicators such as deposit flows, bond yield spreads, and market sentiment shifts.
- Step 2: Forecast outputs dynamically update state variables within an MDP framework representing the bank's funding structure, asset liquidity profile, and market conditions.
- **Step 3:** Forecast probabilities from the ML layer are used to update the transition model, or the belief over next-period states $P(s' \mid s, a)$ rather than overwrite the current state s. Specifically, the one-step forecast $\tilde{p}_{t+1} = \Pr(s_{t+1}^{liq} = 1 \mid x_t)$ parametrizes the action-aware kernel $P_\alpha(s' \mid s, a)$ (see §3.9 - 3.10), which the MDP then uses for policy evaluation. This makes the coupling Bayesian/forward-looking: the observed s_t remains in the current state, while the forecast informs the distribution of s_{t+1} conditional on action.
- Step 4: The MDP agent selects optimal actions, such as early refinancing, reserve accumulation, or balance sheet reallocation, based on maximizing a long-term reward function incorporating funding cost, liquidity risk, and regulatory compliance metrics.
- Step 5: As new data becomes available, ML forecasts and MDP policies are adaptively refined to ensure real-time responsiveness.

Operationally, this hybrid system would enable banks to move from reactive liquidity management toward proactive, scenario-driven optimization that better reflects refinancing pressures and real-world frictions.

2.6.2 Empirical Evidence Supporting Hybrid Approaches

Recent works demonstrate the advantages of combining ML with MDP frameworks for financial decision making.

- Wekwete et al. (2023) proposed a Deep ALM architecture that merged neural network-driven forecasts with reinforcement learning to dynamically rebalance funding strategies.
- Antony and Kumar (2024) evaluate ML models KNN, SVM, Decision Tree, RF, and XGBoost for forecasting liquidity risk in Indian banks, finding KNN and XGBoost most effective for early warning system development.
- Neagu et al. (2024) modeled price movements in illiquid markets using LSTM, while a RL agent determined the optimal timing and scale of hedging actions.

In representative studies, neural sequence models or tree ensembles produce risk or state forecasts that are then consumed by an MDP/DRL layer to select actions under an explicit reward and constraint set. For example, LSTM-based volatility forecasters have been embedded in DRL agents to manage multi-asset portfolios; boosted-tree probability forecasts have been injected into MDP to guide liquidity interventions; and deep ALM architectures have merged network-based forecasts with reinforcement learning to rebalance funding dynamically. Related studies model price moves in illiquid markets with sequence models while a reinforcement agent times hedging actions. Across these settings, integrated ML+MDP systems tend to outperform single-method baselines on

decision-relevant metrics, with ML improving state recognition and risk estimation and the MDP/DRL layer aligning actions to long-horizon objectives.

At the same time, operationalizing the ML to MDP integration remains challenging. Key technical barriers include: (i) model drift, where forecast distributions change as regimes change; (ii) reward specification, which must include tradeoffs between liquidity protection, cost of funds, regulatory penalties; and (iii) exploration - exploitation in DRL, which entails appropriate tuning and offline evaluation in order to prevent inappropriate risk-taking during periods of stress. These integration issues necessitate systems design which permits retraining with frequency, drift monitoring on a continuous basis, reproducible pipelines and auditable change control consistent with Supervisory Letter SR 11-7 on model risk management (Federal Reserve, 2011).

2.6.3 Application of Hybrid Frameworks in Regional Banks

Although leading financial institutions have begun adopting hybrid frameworks that blend machine learning with decision-making models, regional banks in the U.S., despite being among the most exposed to liquidity pressures, have largely remained on the sidelines. These institutions often face clustered refinancing schedules, limited access to diversified funding sources, and operate in environments with less robust data infrastructure.

A well-designed hybrid model can help address these challenges by supporting:

- Strategic management of liquidity buffers in response to evolving macroeconomic trends
- Planning for refinancing events, particularly involving callable or maturing debt instruments

- Prioritization of market-based funding options using sentiment analysis and spread projections

To ensure practical implementation, these systems must meet key operational and regulatory standards by offering:

- Transparency, using tools like SHAP values (Lundberg & Lee, 2017) or rule-based explanations to support model auditability
- Regulatory alignment, meeting supervisory expectations outlined in Office of the Comptroller of the Currency (OCC) Bulletin 2011-12 guidance on model governance and risk management

A practical illustration is provided by Deep Treasury Management models (Englisch et al., 2023), which integrate yield curve simulations, refinancing dynamics, and liquidity constraints into hybrid optimization systems. The Englisch et al. illustration reflects large-institution practice or academic prototypes; we are not aware of peer-reviewed evidence of implementation at U.S. regional banks.

2.7 Literature Gaps and the Timeliness of this Praxis

Although academic research on hybrid MDP - ML frameworks has advanced significantly, key gaps persist, particularly in their application to liquidity management for regional banks. These gaps highlight the need for further exploration and practical adaptation of such models to address the unique challenges faced by these institutions.

2.7.1 Absence of Refinancing-Cycle-Awareness Frameworks

Here, “conventional modeling approaches” refers broadly to traditional ALM techniques and common ML/RL implementations that abstract liquidity as a smooth process and omit refinancing-cycle structure. Many existing ML and RL models

approach liquidity risk as a smooth, continuous phenomenon, overlooking structural elements that shape real-world outcomes, such as refinancing walls, callable debt concentration, and uninsured deposit exposure. These same features are also typically absent from conventional modeling approaches as defined here. The omission weakens stress identification and timing, particularly around refinancing windows.

Such refinancing-specific dynamics are particularly relevant for regional banks, which often face concentrated exposures and limited flexibility. Yet, these factors are frequently absent from conventional modeling approaches (Jiang et al., 2023).

2.7.2 Underutilization of Macroeconomic and Sentiment Signals

Natural language processing (NLP) techniques applied to sources like earnings call transcripts, regulatory statements, and bondholder updates have shown meaningful predictive power in asset pricing contexts (Nassirtoussi et al., 2014). While liquidity is fundamentally a function of cash flows and funding costs, investor sentiment and forward-looking market signals have proven highly predictive under crisis conditions yet remain underutilized.

2.7.3 Limited Model Interpretability for Supervisory Use

While deep learning and DRL techniques provide considerable modeling power, they have been criticized for a lack of transparency. For controlled uses such as liquidity planning, regulators expect models to be subject to audit, explanation and governance throughout their life cycle. The regulatory standard applicable in U.S. banking is SR 11-7/OCC 2011-12, which calls for the documentation of the model purpose, data lineage, assumptions, limitations, evidence of validations, and monitoring, and the use of techniques that facilitate effective challenge (Office of the Comptroller of the Currency,

2011). For contemporary machine learning models, tools such as SHAP provide case-level and global explanations that can be mapped to the documentation and monitoring requirements under SR 11-7 (Lundberg & Lee, 2017). In addition, complementary risk-management practices of NIST AI RMF 1.0 place additional emphasis on transparency, measurements, and governance controls that are consonant with the standards of bank model risk (National Institute of Standards and Technology, 2023).

2.7.4 Inaccessibility for Mid-Sized Institutions

Many existing hybrid models are overly complex or resource-intensive, making them impractical for regional banks that operate with constrained data systems and limited in-house analytical expertise. What is needed are streamlined, scalable solutions built on clear, transparent assumption and models that can be incrementally adopted by treasury teams without requiring significant technical overhaul or specialized staffing.

2.8 Summary and Transition to Methodology

This literature review examines the progress related to the modeling of liquidity risk surrounding U.S. regional banks, including theoretical bases, empirical applications, and practical implementation. It highlights the drawbacks of traditional ALM models, the power of the forecasting but static capabilities of current ML models, and the power of MDP for dynamic analysis and decision making.

The combination of ML forecasting models with MDP-based optimization creates a forward-looking, proactive liquidity management approach. This combination allows for the identification of emerging risks and enables forward-looking and proactive planning.

The proactive approach is especially useful in periods of refinancing when liquidity challenges arise and the need for well-timed interventions is most pronounced.

This praxis develops and evaluates a hybrid model that integrates predictive algorithms, trained on market signals, macroeconomic indicators, and sentiment data, with an MDP-based decision agent whose environment is constructed from bootstrapped liquidity stress scenarios derived from historical refinancing events. The methodological framework will be detailed in Chapter 3 with data inputs, model architecture and construction, training procedures and evaluation metrics.

Chapter 3—Methodology

3.1 Introduction to Methodological Framework

This chapter specifies the end-to-end methodology that transforms raw bank-quarter data into testable outputs and decision evaluations. This process has three stages. Stage 1 defines the data foundation, variable construction, and labeling procedures that produce the modeling matrix (§§3.1- 3.3). Stage 2 specifies the predictive models: tree ensembles and sequence models, and documents training, validation, diagnostics, and robustness (§§3.4–3.8). Stage 3 builds the decision framework: a Markov Decision Process (MDP) on the same state space, an MDP-augmented feature set for prediction, and a Deep Q-Network (DQN) policy that converts model outputs into actions and evaluates them over a four-quarter horizon H4 (§§3.9–3.12). Chapter 4 then reports the corresponding results in the same order: descriptive statistics, predictive performance, state dynamics, scenarios, monitoring, and policy outputs.

This study advances liquidity risk modeling with a hybrid MDP–ML framework for forecasting and optimizing liquidity around refinancing cycles. The architecture distinguishes three roles: (i) an MDP-style state representation whose transition kernel is estimated empirically from panel data (with optional supervised refinement); (ii) predictive learners (RF, XGBoost, LSTM) that consume this state representation and other features to forecast liquidity outcomes; and (iii) policy learning via DQN, which takes the empirically estimated transition kernel and a reward function to learn action policies. Reinforcement learning is applied only for policy optimization; it is not used to estimate transition probabilities.

The objective is twofold. The first is (1) to statistically and operationally predict the probability and severity of liquidity shortfalls in such a way that the estimates can be made under dynamic market and balance sheet conditions and (2) to convert the results of these investigations into procedures which are sufficiently pertinent to give rise to a policy aligned with the Asset - Liability Committee (ALCO) decision process. In so doing, the framework addresses the central regulatory expectations relating to model transparency, model stability and supervisory consistency which is articulated in SR 11-7 and OCC 2011-12.

The methodological innovation lies in encoding liquidity path dependence, how past liquidity states influence future outcomes, via an empirically estimated MDP transition kernel. These state variables (PrevState, P_high_next_from_prev) are embedded within the ML layer, transforming static forecasts into state-aware, sequential predictions. Each algorithm is learnt from a unified, multi-domain dataset combining Trade Reporting and Compliance Engine (TRACE) bond microstructure, Call Report fundamentals, and systemic volatility measures (Chicago Board of Options Exchange CBOE Volatility Index VIX and yield volatility).

This hybrid structure creates a closed feedback loop between predicting and then optimizing decisions. The ML layer predicts conditional probability distributions of the realms of liquidity, and the MDP layer finds optimal liquidity preserving actions given the existing regime. Together, this gives a quantitative, regulator-compliant architecture with predictive power, interpretability and real time application, such that bank treasurers and their teams can best assess funding stress, calibrate their liquidity buffers, and

provide the markets effectively with funding and confidence with both accuracy and foresight.

One-way coupling and kernel scope: during supervised learning, the MDP contributes only historical, unconditional state persistence statistics that are computed before model training and kept fixed. Specifically, this praxis estimates a two-state transition matrix $\hat{P}^{\text{hist}}$ from observed liquidity states $s \in \{0,1\}$ over the training window, independent of actions. From this, the praxis forms the lagged covariates $\text{PrevState}_{i,t} = s_{i,t-1}$and $P_{i,t}^{\text{high,hist}} = \hat{P}^{\text{hist}}(s_{i,t-1} \to 1)$.

- $\text{PrevState}_{i,t} = s_{i,t-1}$ is a lagged liquidity indicator showing whether bank *i* was in stress (1) or normal (0) in the previous period, which captures state persistence (if a bank was stressed last quarter, it might remain stressed).
- $P_{i,t}^{\text{high,hist}} = \hat{P}^{\text{hist}}(s_{i,t-1} \to 1)$ is the historical probability of entering a high-liquidity-risk state given the previous state. For each bank *i* and time *t*, it tells the model how likely stress is to continue or emerge given the most recent state.

These features are passed to the ML models as ordinary predictors. No ML output is used to build $\hat{P}^{\text{hist}}$, and the kernel is not updated by forecasts in the supervised stage. The action-aware mixture $P_\alpha(s' \mid s, a)$ appears only in the decision layer (§§3.10–3.11) after a forecaster has been selected and frozen.

3.2 Data Sources and Integration

The research employs a multi-domain dataset integrating market microstructure, regulatory financial statements, and macro-financial volatility indicators to capture both cross-sectional and temporal dynamics of bank liquidity. The unified panel is constructed

at the bank - quarter frequency for all U.S. regional banks active between 2013 and 2023, ensuring consistency across regulatory reporting cycles and refinancing events.

The first dataset, the Trade Reporting and Compliance Engine (TRACE) dataset, provides transaction-level data on secondary-market trades of bank-issued debt securities. From TRACE, the study extracts trade volume, execution price, and transaction count. These are aggregated to the quarterly level to form microstructure-based liquidity measures that reflect market depth and stability. The second dataset comprises Call Report filings (FFIEC 031/041), including balance sheet, income, and capital items such as total assets (ATQ), total equity (SEQO), long-term debt (DLTTQ), short-term debt (DLCQ), and depreciation (DPQ). These variables characterize each bank's funding structure, leverage, and capital adequacy. The dataset is accessed via WRDS Bank Fundamentals Quarterly, which integrates FFIEC 031/041 and FR Y-9C filings while preserving Compustat-style variable names for consistency across historical periods. Thus, codes such as ATQ and DLTTQ correspond to FFIEC variables (e.g., RCFD2170 for total assets) through WRDS's standardized mapping. These variables characterize a bank's funding structure, leverage, and capital adequacy.

To capture systemic and market-wide stress, the third dataset introduces macro-volatility indicators, specifically the Chicago Board Options Exchange Volatility Index (VIX) and a derived bond yield volatility measure. Both are averaged quarterly and aligned to the same temporal index, quarter_key. This ensures synchronization between market-level and institution-level observations.

All data sets are preprocessed in a standardized pipeline that converts financial reporting dates into quarters, harmonizes ticker symbols and identifiers, and applies strict

controls for data quality. Variables with greater than 50 percent missingness were dropped. Remaining variables with missingness were imputed using medians for the specific variable. The numerical variables were standardized using z-scores to allow for comparability across banks. The resulting dataset forms a complete, modeling-ready panel that allows the hybrid MDP - ML framework to learn liquidity behavior across diverse banks and economic regimes with statistical precision and regulatory interpretability.

3.3 Feature Engineering and Label Construction

This section defines the liquidity constructs, engineered predictors, regime indicators, and targets used by the hybrid MDP - ML framework. All variables are aligned at the bank-quarter level and use information available at time to only.

3.3.1 Market Microstructure Liquidity

Let $V_{i,t}$be total traded par volume for bank i in quarter t; $N_{i,t}$the number of executions; and $\{P_{i,t,k}\}_{k=1}^{N_{i,t}}$the execution prices.

Define the quarter-level price dispersion as the sample standard deviation of prices: $\bar{P}_{i,t} = \frac{1}{N_{i,t}} \sum_{k=1}^{N_{i,t}} P_{i,t,k}$ as the average trade price, and

$\text{PriceStd}_{i,t} = \sqrt{\frac{1}{N_{i,t}-1} \sum_{k=1}^{N_{i,t}} (P_{i,t,k} - \bar{P}_{i,t})^2}$ as the price standard deviation, a measure of price dispersion or volatility of execution prices across all trades that quarter. A higher PriceStd means greater disagreement or illiquidity in market pricing (trades occur at more dispersed prices).

When daily returns and volumes are available, we compute Amihud illiquidity; where price updates are sufficiently frequent, we compute Roll's effective spread. If a

measure has >50% missingness in a bank-quarter it is dropped; otherwise, missing entries are mean-imputed as in §3.2, so LiqScore remains fully defined.

$\text{ILLIQ}_{i,t} = \frac{1}{D_{i,t}} \sum_{d=1}^{D_{i,t}} \frac{|R_{i,t,d}|}{\text{Vol}_{i,t,d}}$,where $R_{i,t,d}$is the daily return and $Vol_{i,t,d}$the daily trading volume. This captures how much the price moves per unit of trading volume or how sensitive prices are to trading.

Where price updates are sufficiently frequent, Roll's effective spread estimator (Roll, 1984) $\widehat{\text{Roll}}_{i,t} = 2\sqrt{\max\{0, -\widehat{\text{Cov}}(\Delta P_{i,t,d}, \Delta P_{i,t,d-1})\}}$, inferring the effective bid-ask spread from negative serial covariance of price changes.

Here, the Amihud illiquidity measure quantifies how much a security's price moves in response to a given amount of trading volume, capturing the price impact of the bond trades. It is calculated as the average ratio of the absolute daily return to the trading volume, with higher values indicating lower liquidity and greater price sensitivity to trades (Amihud, 2002). Roll's effective spread estimator measures transaction costs by inferring the bid - ask spread from the negative serial covariance of successive price changes. Larger Roll means wider spreads and lower liquidity.

Now we can construct a composite liquidity index using robust z-scores. For any variable x, $z_r(x) = (x - \text{med}(x))/(1.4826\,\text{MAD}(x))$, with $\text{med}(x)$ as the sample median and Median Absolute Deviation (MAD) $\text{MAD}(x) = \text{med}(|\, x - \text{med}(x)\,|)$. The factor 1.4826 makes MAD consistent with the normal standard deviation. MAD and med computed once on the training window and the held fixed. MAD measures spread around the median and therefore is robust to outliers as compared to standard deviation. Inputs are winsorized at the 1st/99th percentiles and lagged one quarter to preclude look-ahead.

3.3.1.1 Composite Liquidity Score

Define the sign-aligned, robust-standardized vector $\tilde{Z}_{i,t} = [z_r(V_{i,t}),\ z_r(N_{i,t}),\ -z_r(\text{PriceStd}_{i,t}),\ -z_r(\text{ILLIQ}_{i,t})]$, with $z_r(V_{i,t})$ being robust z of traded volume, $z_r(N_{i,t})$ being robust z of number of executions, $-z_r(\text{PriceStd}_{i,t})$ being robust z of price dispersion, flipped so higher score would represent more liquidity, and $-z_r(\text{ILLIQ}_{i,t})$ being robust z of Amihud illiquidity, flipped for the same rationale. Altogether, larger components uniformly indicate more liquidity.

Unit-weight index: $\text{LiqScore}_{i,t}^{\text{unit}} = z_r(V_{i,t}) + z_r(N_{i,t}) - z_r(\text{PriceStd}_{i,t}) - z_r(\text{ILLIQ}_{i,t})$.

For the Principal Component Analysis[2] (PCA) - weighted index: we will estimate PCA on the training-sample correlation matrix of $\tilde{Z}$.

Let $\hat{w}$ be the unit-length loading vector of PC1; flip its sign as needed so that $\hat{w}_1 + \hat{w}_2 > 0$ (positive combined loading on $V$and N).

We can score any period by $\text{LiqScore}_{i,t} = \hat{w}^{\top}\tilde{Z}_{i,t}$

Here neither $\hat{w}$ nor the robust scaling statistics are refit on validation or test sets.

Binary state (contemporaneous, cross-sectional):

$\text{HighLiquidity}_{i,t} = 1\{\text{LiqScore}_{i,t} > \text{median}_{\,j}(\text{LiqScore}_{j,t})\}$ and ties at the median default to 0 (or a deterministic tie-break, e.g., higher volume).

3.3.2 Balance Sheet and Capital Structure

To characterize the liquidity and solvency profile of each bank, we can derive a set of balance sheet and capital structure ratios from U.S. Call Report data. Let the Loan-to-Deposit Ratio (LDR), Common Equity Tier 1 Ratio (CET1), Risk-Weighted Assets

[2] (Jolliffe, 2002)

(RWA), High-Quality Liquid Assets (HQLA), Short-Term Funding Ratio (ST Funding), and Short-Term Debt (DLCQ) be defined as follows:

$$\mathrm{LDR}_{i,t} = \frac{\mathrm{Loans}_{i,t}}{\mathrm{Deposits}_{i,t}}, \mathrm{CET1}_{i,t} = \frac{\mathrm{CET1\_Capital}_{i,t}}{\mathrm{RWA}_{i,t}}$$

$$HQLA_{i,t} \approx \frac{\mathrm{Cash}_{i,t}+\mathrm{UST}_{i,t}+\mathrm{Agency}_{i,t}}{\mathrm{Assets}_{i,t}}, ST\ Funding_{i,t} = \frac{DLCQ_{i,t}}{Liabilities_{i,t}}$$

with LDR captures funding tightness (higher ⇒greater reliance on deposits to fund loans), CET1 measures loss-absorbing capital relative to risk-weighted assets, HQLA is proxied by cash plus high-grade sovereign/agency securities (consistent with Basel LCR logic) scaled by total assets, and ST Funding quantifies rollover exposure risk via short-term debt (DLCQ) as a share of total liabilities. These ratios jointly capture the interaction between funding structure, capital adequacy, and liquidity resilience, which are core determinants of a bank's capacity to absorb shocks and sustain operations under stress.

Funding concentration: For funding categories $f \in F$, we can calculate the Herfindahl index[3] as $\mathrm{HHI}_{i,t}^{\mathrm{fund}} = \sum_{f\in F} s_{i,t}(f)^2, s_{i,t}(f) = \frac{\mathrm{Funding}_{i,t}(f)}{\sum_{g\in F}\mathrm{Funding}_{i,t}(g)}$, with F representing the set of all funding categories (e.g., deposits, repos, commercial paper, senior debt, subordinated debt, etc.), $f \in F$as a specific funding source type (e.g., "wholesale deposits" or "term repo"), $\mathrm{Funding}_{i,t}(f)$ as the amount of funding bank i raises from source $f$in quarter t, $s_{i,t}(f)$ as the share (or weight) of funding source $f$in the bank's total funding mix.

The Refinancing pressure (default horizon $K = 4$ quarters, defining the near-term "maturity wall.") can be defined as $\mathrm{RefiLoad}_{i,t} = \frac{\text{Debt maturing in } [t,t+K]}{\text{Total wholesale debt}_{i,t}}$ where the

[3] (Herfindahl, 1950)

numerator represents the total face value of outstanding wholesale debt instruments (e.g., senior notes, covered bonds, CDs, or term repos) that come due between quarter t and quarter $t + K$ (the refinancing burden within the next Kquarters) , and the denominator is the total outstanding wholesale funding for bank i at time t, including all market-based borrowings excluding retail deposits (e.g., interbank, repo, commercial paper, medium-term notes, covered bonds, and senior unsecured debt) and serving as a scaling base to express the maturity load as a fraction of total market funding exposure. Here the ratio $\text{RefiLoad}_{i,t}$ represents a refinancing intensity metric, showing how much of a bank's wholesale funding will mature soon: High values mean heavy near-term refinancing needs and higher rollover risk and vice versa.

3.3.3 Macro and Volatility Factors

To capture systemic uncertainty and macro-financial conditions, a set of quarterly market factors is constructed from high-frequency data:

VIX_t = quarterly mean of daily implied volatility (VIX).

$\text{YldVol}_t = \frac{1}{D_t - 1} \sum_{d=1}^{D_t} (\Delta y_{t,d} - \Delta \bar{y}_t)^2$, where $\Delta y_{t,d}$denotes the daily change in the 10-year Treasury yield and $D_t$the number of trading days in quarter t. YldVol_t thus measures the realized yield volatility within the quarter.

The term-structure slope is defined as $\text{Slope}_t = y_t(10\text{y}) - y_t(2\text{y})$ providing a gauge of monetary-policy stance and future growth expectations.

When available, a credit-spread proxy from Trade Reporting and Compliance Engine[4] (TRACE) transaction data, typically the median secondary-market spread between BBB and A-rated financial bonds, is included to capture funding-market risk premia.

[4] (Financial Industry Regulatory Authority)

In combination, $\{VIX_t, \text{YldVol}_t, \text{Slope}_t, \text{CreditSpread}_t\}$ provide a compact, orthogonal summary of the macroeconomic volatility, policy stance, and credit risk environment influencing banks' liquidity and refinancing behavior in each quarter.

3.3.4 Regimes and Event Time

Refinance (issuance/maturity pressure):

The refinancing-pressure indicator, RefiWindow, flags bank–quarters with elevated maturity concentration. It is defined as $\text{RefiWindow}_{i,t} = \mathbb{1}\,(V_{i,t} \geq \tau_V)$, where $V_{i,t}$ denotes the maturity volume ratio, and τ_V is a threshold (set to the 75th percentile of the sample distribution of $V_{i,t}$) that marks unusually high refinancing exposure.

- The maturity volume ratio is computed as $V_{i,t} = \frac{\text{MaturingDebt}_{i,t}}{\text{TotalLiabilities}_{i,t}}$, where $\text{MaturingDebt}_{i,t}$ includes all bond, note, and term-deposit obligations with residual maturity ≤ 1 year (proxied using Call Report schedule RC–M and RC–Q items for time deposits and long-term borrowings), and
- $\text{TotalLiabilities}_{i,t}$ corresponds to FFIEC variable RCON2948.

For comparability, both numerator and denominator are averaged within quarter t. RefiWindow thus equals 1 when a bank's short-term funding load exceeds the 75th percentile of its peer distribution, signaling a refinancing window with elevated rollover risk.

$\text{RefiWindow}_{i,t} = 1\{\text{RefiLoad}_{i,t} \geq \tau_V \text{ or } V_{i,t} \geq \tau_{V,75}\}$, where $\text{RefiWindow}_{i,t}$ indicates that bank i in quarter t faces elevated refinancing pressure, $\text{RefiLoad}_{i,t}$ measures the ratio of maturing debt (or issuance load) relative to total liabilities, $V_{i,t}$denotes issuance or maturity volume for bank $i$in quarter t, and $\tau_{V,75}$ represents the cross-sectional 75th percentile of issuance or maturity volume at time t. This binary indicator captures quarters where refinancing intensity or maturity concentration exceeds systemic norms.

Market Stress Regime

$\text{StressRegime}_t = 1\{\text{VIX}_t > \bar{\text{VIX}} + 1\sigma_{\text{VIX}} \text{ or } \text{YldVol}_t > \bar{\text{YldVol}} + 1\sigma_{\text{YldVol}}\}$, where: VIX_t is the CBOE Volatility Index[5] (proxy for market-wide risk aversion), YldVol_t represents the yield volatility of benchmark Treasury rates or credit spreads, and $\bar{\text{VIX}}$, σ_{VIX}, $\bar{\text{YldVol}}$, σ_{YldVol} represent the mean and standard deviation of their respective time series. The indicator equals one in quarters where volatility or yield dispersion exceeds one standard deviation above its historical mean, identifying systemic stress periods.

Event Time Alignment

Each bank-quarter observation is re-centered around the quarter of maximum refinancing pressure: $k \in [-K, +K]$,with$k = 0$ denoting the refinancing peak.

This event-time alignment enables comparative analysis of liquidity, spreads, and balance-sheet behavior before and after refinancing stress events.

3.3.5 MDP-Derived State Features

Empirical Transition Kernel (Dirichlet-Smoothed):

[5] (Cboe Global Markets)

Let the discrete liquidity state for bank $i$in quarter t be $s_{i,t} \in \{0,1\}$, where 0 represents a low-liquidity regime and 1 represents a high-liquidity regime. Over the training sample, let $n(u \to v)$ represent the number of observed transitions from state u at time $t-1$ to state v at time t. The empirical transition kernel, smoothed with a symmetric Dirichlet prior of strength $\alpha > 0$, is estimated as

$$\hat{P}_{\text{hist}}(u \to v) = \frac{n(u \to v) + \alpha}{\sum_{w \in \{0,1\}} n(u \to w) + 2\alpha}, \alpha > 0$$

This smoothed kernel mitigates sampling noise when the number of transitions is small, while ensuring that all state transitions retain non-zero probability mass.

Derived Markovian Features

From the estimated kernel, two lagged features are constructed for each bank-quarter observation: $\text{PrevState}_{i,t} = s_{i,t-1}, P_{i,t}^{\text{high,hist}} = \hat{P}_{\text{hist}}(s_{i,t-1} \to 1)$, where $\text{PrevState}_{i,t}$ captures the first-order Markov dependency by recording the prior period's liquidity regime, $P_{i,t}^{\text{high,hist}}$represents the empirical transition probability of moving into a high-liquidity state, conditional on the bank's previous state.

Estimation and Information Set

Transition probabilities can be estimated under several pooling strategies: 1) Pooled across all banks to obtain a single aggregate kernel; 2) Regime-conditioned (e.g., computed separately under high-stress versus low-stress market regimes), or 3) Bank-clustered (estimated within peer groups or size deciles).

Regardless of specification, the estimation uses information available at time $t$or earlier only, thereby maintaining strict temporal causality. At prediction time for quarter $t+1$, the feature $P_{i,t}^{\text{high,hist}}$ is computed solely from the realized state $s_{i,t}$, ensuring a fully

feed-forward, leakage-free representation consistent with out-of-sample forecasting standards.

3.3.6 Temporal Derivatives and Interactions

Quarter-over-Quarter Differences and Multi-Quarter Slopes

To capture short-term dynamics and directional trends in each predictor $x_{i,t}$, quarter-over-quarter changes and L-quarter rolling slopes are computed as follows:

$\Delta x_{i,t} = x_{i,t} - x_{i,t-1}, \text{Slope}_L(x)_{i,t} = \frac{1}{L}\sum_{l=1}^{L}(x_{i,t-l+1} - x_{i,t-l})$, where $\Delta x_{i,t}$measures the most recent quarter-to-quarter change in a variable, while $\text{Slope}_L(x)_{i,t}$ represents the average directional momentum of that variable across the prior L quarters.

These temporal derivatives highlight acceleration or deceleration in liquidity, funding, and pricing metrics, allowing the model to detect emerging liquidity pressures before they become observable in levels alone.

State-Dependent Interactions

To account for nonlinear amplification effects under stress or refinancing pressure, a set of multiplicative interaction features is constructed. These capture how market and balance-sheet sensitivities vary across regimes:

$\text{RefiWindow}_{i,t} \times \text{YldVol}_t, \text{RefiWindow}_{i,t} \times \text{PriceStd}_{i,t}, \text{StressRegime}_t \times \text{STFunding}_{i,t}$, where $\text{RefiWindow}_{i,t} \times \text{YldVol}_t$ measures the joint effect of refinancing concentration and yield volatility on liquidity risk, $\text{RefiWindow}_{i,t} \times \text{PriceStd}_{i,t}$ captures the interaction between refinancing pressure and asset-price dispersion, linking market depth with liability maturity concentration, $\text{StressRegime}_t \times \text{STFunding}_{i,t}$ reflects sensitivity of short-term funding reliance during periods of heightened market stress.

These cross-terms introduce conditional heterogeneity into the feature space, enabling the machine-learning component of the hybrid framework to approximate regime-specific behaviors and nonlinear propagation of liquidity shocks.

3.3.7 Targets for Classification and Severity

Liquidity Score

For a forecasting horizon of $k$quarters ahead and a cross-sectional lower-tail quantile threshold q_τ(e.g., $\tau = 0.40$), define the binary target on liquidity score $\text{LiqScore}_{i,t}$: $y_{i,t}^{(k)} = 1\left\{\text{LiqScore}_{i,t+k} < q_\tau\,,\left(\text{LiqScore}_{\cdot,\,t+k}\right)\right\}$, where $q_\tau(\text{LiqScore}_{\cdot,\,t+k})$ is the τ-quantile of the cross-section of banks at time $t + k$.

Here, we have the liquidity score derived from multiple microstructures (3.3.1 above) and balance sheet features, normalized within each quarter to ensure cross-bank comparability. The score summarizes a bank's ability to meet short-term obligations without incurring significant funding costs or asset sales. A lower value implies tighter liquidity and higher vulnerability.

Lower-Tail Classification Target

Lower-Tail Classification Target $y_{i,t}^{(k)}$ labels future liquidity shortfalls as binary events. $y_{i,t}^{(k)} = 1$ indicates that bank i's future liquidity score falls in the lowest τfraction of the cross-sectional distribution at $t + k$, representing relative underperformance or stress compared to peers.

Severity Target

Putting together, the following severity target $y_{i,t}^{\text{sev},(k)}$ measures how far below the critical liquidity threshold a bank falls and is defined as:

$y_{i,t}^{\text{sev},(k)} = \max\{0,\ q_\tau\ (\text{LiqScore}_{\cdot,\,t+k}) - \text{LiqScore}_{i,t+k}\}$, where a value of zero means the bank's liquidity is adequate, while positive values indicate deficit severity. The severity target is a continuous measure that can support regression modeling, capturing how deep a bank's liquidity deterioration is relative to systemic norms. $y_{i,t}^{\text{sev},(k)}$ is especially useful for MDP reward shaping and calibration of liquidity buffer recommendations.

Liquidity Coverage Ratio (LCR)-oriented target

To ensure alignment with regulatory liquidity metrics, a target based on the Liquidity Coverage Ratio (LCR) can be defined as

$$y_{i,t}^{\text{LCR\_gap},(k)} = \max\{0,\ 1.00 - \text{LCR}_{i,t+k}\}, \text{ with } \text{LCR}_{i,t} \approx \frac{\text{HQLA}_{i,t}}{\text{NetOutflows}_{i,t}}$$

where $y_{i,t}^{\text{LCR\_gap},(k)}$ measures the forward shortfall relative to the regulatory minimum LCR requirement of 1.00, capturing the extent to which a bank's projected high-quality liquid assets (HQLA) fail to cover expected net outflows under stress. This metric can help link model predictions to supervisory benchmarks such as Basel III liquidity standards, translating abstract liquidity risk into policy-relevant measures.

In this praxis, $\text{NetOutflows}_{i,t}$ is proxied from Call Report data following the supervisory construction in Basel III and U.S. LCR Rule (12 CFR 249). For each bank $i$and quarter t, we approximate:

$$\text{NetOutflows}_{i,t} \approx 0.10 \times \text{UninsuredDeposits}_{i,t} + 0.25 \times \text{WholesaleFunding}_{i,t} + 0.05 \times \text{StableRetailDeposits}_{i,t} - 0.50 \times \text{AvailableCredit}_{i,t}.$$

The coefficients represent assumed runoff and inflow rates drawn from Basel liquidity outflow tables, scaled to the reporting granularity of FFIEC 031/041 variables. When Call Report details were insufficient, category balances were imputed using corresponding H8 or FDIC Summary of Deposits proportions. This construction yields a reproducible proxy for stressed net outflows while maintaining consistency across banks and time.

3.3.8 Normalization, Leakage Control, and Drift

Robust scaling and standardization

All continuous predictors are transformed using a robust z-score computed only on the training data and then held fixed for validation and test periods.

Let $x_{i,t}$ denote a raw feature and let $\tilde{x}$ be the median on the training fold, with median absolute deviation $\mathrm{MAD}(x) = \mathrm{median}(\ |\ x - \tilde{x}\ |)$, we can define the robust standardized value $z_r(x_{i,t}) = \frac{x_{i,t} - \tilde{x}}{\mathrm{MAD}(x)} \times 1.4826$ where factor 1.4826 yields a consistent estimator of the standard deviation under normality.

The operator $z_r(\cdot)$ is computed once per feature on the training fold and applied unchanged to validation and test sets. To preserve temporal causality, all predictors are lagged by at least one quarter before scaling and modeling.

Sequence models and windowing

For sequence learners, features are assembled in fixed-length windows $X_{i,t-L+1:t} = [x_{i,t-L+1}, x_{i,t-L+2}, \ldots, x_{i,t}]$, with the prediction target at $t + 1$. Padding and masks are used when the entity history is shorter than L. All elements in $X_{i,t-L+1:t}$ are constructed from data available at or before t, scaled with the frozen $z_r(\cdot)$, and lagged to

prevent contemporaneous leakage. Sliding windows advance with stride one unless otherwise specified.

Drift monitoring and stability tests

To monitor covariate and score stability between the development (train-time reference) and production (recent scoring window) populations, the Population Stability Index (PSI) is computed on key features and model scores.

Partition the range of a variable into B bins. Let $p_b$and $q_b$be the proportions in bin $b$for development and production respectively. The PSI is

PSI $= \sum_{b=1}^{B}(p_b - q_b)\ \ln\left(\frac{p_b}{q_b}\right)$, here we will assign thresholds PSI ≤ 0.10 to indicate immaterial shift, $0.10 <$ PSI ≤ 0.25 to indicate moderate shift, and PSI > 0.10 to trigger a model review in this study. Values > 0.25 often warrant recalibration or retraining.

The engineered matrix, labels, and MDP state features provide a leak-resistant, regulator-aligned basis for state-aware learning and policy evaluation (Pástor and Stambaugh, 2003; Puterman, 1994).

3.4 Model Classes and Training Protocol

Objective

For each bank–quarter i, t, the praxis learns calibrated probabilities $p_{i,t+1} = \Pr\left(s_{i,t+1} = 1 \mid x_{i,t}\right)$ using features observable at t. Let $\sigma(z) = (1 + e^{-z})^{-1}$

Primary loss (class-weighted logistic)

$\mathcal{L}(\theta) = -\sum_{(i,t)\in\mathcal{D}} w_{i,t}\,[y_{i,t+1}\log\, p_{i,t+1} + (1 - y_{i,t+1})\log\,(1 - p_{i,t+1})], p_{i,t+1} = \sigma(f_\theta(x_{i,t}))$, with $w_{i,t} \propto 1/\Pr\,(y)$to correct imbalance.

Model classes

Random Forest (RF)

An ensemble of T bootstrapped, feature-subsampled trees with Gini split criterion

$\text{Gini}(S) = 1 - \sum_{c\in\{0,1\}} \pi_c^2 , \pi_c = \frac{1}{|S|} \sum_{(x,y)\in S} \mathbf{1}\{y = c\}$

Prediction $p_{\text{RF}}(x) = \frac{1}{T} \sum_{t=1}^{T} h_t(x)$.

Depth, min-leaf, and m_{try}tune variance.

Gradient-boosted trees (XGBoost)

Additive trees $f_k \in \mathcal{F}$with second-order objective

$\mathcal{L}^{(t)} = \sum_i (l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i))) + \Omega(f_t), \Omega(f) = \gamma T + \frac{\lambda}{2} \| w \|_2^2$, with leaf weight

$w_j^* = -\frac{\sum_{i\in I_j} g_i}{\lambda + \sum_{i\in I_j} h_i}$where $g_i = \partial_{\hat{y}} l, h_i = \partial_{\hat{y}}^2 l$.

Long Short-Term Memory

Windows and indices

For each bank and prediction quarter t, define the input window

$\mathcal{W}_t = \{t - L + 1, \dots, t\}$ with feature vectors $\mathbf{x}_\tau \in \mathbb{R}^p$for $\tau \in \mathcal{W}_t$. The model consumes the ordered sequence $(\mathbf{x}_{t-L+1}, \dots, \mathbf{x}_t)$ and outputs probability for the next-quarter state y_{t+1}.

Per-step LSTM recurrence

For each $\tau \in \mathcal{W}_t$the cell updates

$$
\begin{aligned}
\mathrm{i}_\tau &= \sigma(W_i \mathrm{x}_\tau + U_i \mathrm{h}_{\tau-1} + \mathrm{b}_i), \\
\mathrm{f}_\tau &= \sigma(W_f \mathrm{x}_\tau + U_f \mathrm{h}_{\tau-1} + \mathrm{b}_f), \\
\tilde{\mathrm{c}}_\tau &= \tanh(W_c \mathrm{x}_\tau + U_c \mathrm{h}_{\tau-1} + \mathrm{b}_c), \\
\mathrm{c}_\tau &= \mathrm{f}_\tau \odot \mathrm{c}_{\tau-1} + \mathrm{i}_\tau \odot \tilde{\mathrm{c}}_\tau, \\
\mathrm{o}_\tau &= \sigma(W_o \mathrm{x}_\tau + U_o \mathrm{h}_{\tau-1} + \mathrm{b}_o), \\
\mathrm{h}_\tau &= \mathrm{o}_\tau \odot \tanh(\mathrm{c}_\tau),
\end{aligned}
$$

with initial states $\mathrm{h}_{t-L} = 0$ and $\mathrm{c}_{t-L} = 0$ (or learned bank-level initializers).

Output layer and target

After processing the full window, set $\mathrm{h}_t$as the summary representation and map to the one-step-ahead probability $p_{t+1} = \sigma(W_y \mathrm{h}_t + b_y) = \Pr(y_{t+1} = 1 \mid x_{t-L+1:t})$, where $\sigma(\cdot)$ is the logistic link. The training loss is weighted logistic cross-entropy on y_{t+1}.

Variable length panels and masking

If a bank has fewer than L quarters at the start of the panel, this praxis left-pad the sequence with zeros and apply a binary mask so that padded steps do not update gradients. Mini batches are formed by bucketing similar lengths to reduce padding.

Regularization and optimization

Dropout is used on h_τ, recurrent dropout within gates, and L_2 penalty on $W., U.$. Adam optimizes with a plateau scheduler. Early stopping monitors validation AUC.

Two variants are trained: LSTM (No MDP) excludes PrevStateand $P^{\text{high,hist}}$. LSTM (+MDP) includes both features at each τas part of x_τ.

3.5 Calibration, thresholds, and validation splits

Temporal validation

Rolling-origin folds $(\mathcal{T}_r, \mathcal{V}_r)$with $\max(\mathcal{T}_r) < \min(\mathcal{V}_r)$. Hyperparameters chosen by mean validation AUC across folds, final model refit on full train.

Calibration

- Platt scaling: fit $p^* = \sigma(a \log \frac{\hat{p}}{1-\hat{p}} + b)$ on validation.
- Isotonic regression: $p^* = \mathrm{Iso}(\hat{p})$when reliability curves are non-monotone.

Action-aware thresholds. For threshold t, $F_1(t) = \frac{2\,\mathrm{Prec}(t)\,\mathrm{Rec}(t)}{\mathrm{Prec}(t)+\mathrm{Rec}(t)}$, $t^\star = \arg\max_t F_1(t)$,

or cost-aligned: $t^* = \arg\max_t E\,[\alpha \cdot \mathrm{TP}(t) - \beta \cdot \mathrm{FP}(t) - \gamma \cdot \mathrm{FN}(t)]$,

mapping $(\alpha, \beta, \gamma)$to one-step reward deltas in §3.10.

3.6 Regularization, hyperparameters, and ablations

Regularization grid

- RF: depth ∈ {4,6,8,10}, min-leaf ∈ {20,50,100}, trees $T \in \{200,500,1000\}$.
- XGB: max_depth {3,4,5}, learning_rate {0.02,0.05,0.1}, subsample/colsample {0.6,0.8,1.0}, $\lambda \in \{0,1,10\}$, $\gamma \in \{0,1,5\}$, trees with early stopping.
- LSTM: hidden units {32,64,128}, layers {1,2}, dropout/recurrent-dropout {0.1,0.2,0.3}, $L \in \{4,6,8\}$, Adam LR {1e-3,5e-4}.

Ablations

1. No-MDP vs +MDP features.
2. No-macro vs macro.
3. Event-time removal.
4. Reward-weight grid for DQN was used later in §3.11.

Leakage guards

All transformations fit on training folds only; scalers and calibrators were frozen before test.

3.7 Diagnostics: Stability, Variance Inflation Factor (VIF), and Drift Readiness:

Variance Inflation Factor (VIF)

For design matrix X, regress each column X_j on X_{-j} and compute $\text{VIF}_j = \frac{1}{1-R_j^2}$,

Flag VIF > 5 for review; in the panel the de-meaned VIFs are within 1 - 3 (see Chapter 4).

Stability across folds

For feature rankings (e.g., XGB gain), we will compute Spearman ρ across folds; flag median $\rho < 0.6$.

For LSTM, monitor weight norms and early-stop variance.

Drift readiness

Pre-register PSI thresholds and triggers: PSI = $\sum_{b=1}^{B} (p_b - q_b) \ln \frac{p_b}{q_b}$, review if PSI > 0.10, plus two-sample KS on key drivers. Define retraining triggers and logging to support Chapter 4.10.

3.8 Explainability Assets (Importance, SHAP):

Global importance

- RF/XGB: split-gain and permutation importance.
- LSTM: permutation on input channels and integrated gradients over L-step windows.

SHAP for tree models

Shapley value $\phi_j$satisfies additivity and consistency; we compute TreeSHAP to obtain: $f(x) - \mathbb{E}[f(X)] = \sum_j \phi_j(x)$. Here the deliverables include summary plots, dependence plots for rate/vol/spread, and interaction with RefiWindow.

3.9 MDP State Space and Transition Estimations

State

At quarter t, the MDP state is $s_t = (s_t^{\text{liq}}, z_t)$, where $z_t \in \{0,1\}^m$is a compact regime vector (e.g., RefiWindow_t, StressRegime_t).

This praxis adopts a 1-based canonical index set $\mathcal{S}_K = \{1, \dots, K\}$, for all liquidity state

specifications. In the baseline binary case $K = 2$, $1 = \text{Low}, 2 = \text{High}$.

When source features encode liquidity as {0,1}, we apply the bijection $\phi(0) = 1$ and $\phi(1) = 2$ before estimation.

For finer granularity, s_t^{liq}can be discretized into $K \geq 2$ quantiles.

Empirical transitions (Dirichlet smoothing)

Let $\mathcal{S} = \{1, \dots, K\}$be the liquidity state set.

Let $n(u \rightarrow v)$denote the observed number of training-window transitions from state $u$at $t$to state $v$at $t + 1$. Row totals are $n(u) = \sum_{w \in \mathcal{S}} n(u \rightarrow w)$,

With a Dirichlet prior $\alpha_u = (\alpha_{u,1}, \dots, \alpha_{u,K})$on each row, the posterior-mean (additive-smoothing) estimator is $\hat{P}_{u,v} = \frac{n(u \rightarrow v) + \alpha_{u,v}}{n(u) + \sum_{w \in \mathcal{S}} \alpha_{u,w}}, u, v \in \mathcal{S}$.

This praxis uses a symmetric prior $\alpha_{u,v} = \alpha_0 / K$so each row remains stochastic for sparse banks or years. Estimation is time-causal: counts use only data up to t; no information from the test window enters $\hat{P}_\alpha$.

Conditional kernels

When sample size permits, transitions can be conditioned on regimes or bank fixed effects: $\hat{P}_\alpha(v \mid u, z), \hat{P}_\alpha(v \mid u, \text{bank} = i)$, constructed by replacing $n(\cdot)$with counts within each stratum and applying the same Dirichlet smoothing. In practice we report the pooled kernel and a regime-specific kernel for (rate-vol high, spread high)vs. complements (Chapter 4).

Transition features fed back to ML

From $\hat{P}_\alpha$we compute (i) the persistence statistic $\hat{P}_\alpha(u \mid u)$, and (ii) the one-step probability of high state given the lagged state: $P_{i,t}^{\text{high}} \equiv \hat{P}_\alpha(1 \mid s_{i,t-1}^{\text{liq}})$

Together with the indicator $\text{PrevState}_{i,t} = s^{\text{liq}}_{i,t-1}$, these constitute the MDP-derived features used by the supervised models (§3.3.5).

In essence, the methodology here uses a one-way coupling where a historical, action-independent transition matrix is estimated to generate features for the ML models. This is a clean and stable approach. This one-way approach was selected to ensure model stability, prevent recursive dependencies during supervised training, and maintain a clear separation between historical dynamics and model forecasts.

3.10 Reward, Actions, and Constraints

Actions

Let $\mathcal{A} = \{\text{Hold,IssueTerm,SellHQLA,DrawRepo}\}$ with state-dependent admissible set $\mathcal{A}(s_t)$(issuance blackouts, HQLA minima, etc.).

One-step reward

For origin state s_t, action a_t, and realized next state s_{t+1},

$$R(s_t, a_t, s_{t+1}) = \alpha_1 \underbrace{\widetilde{LiqScore}\,(s_{t+1})}_{future\ liquidity} - \underbrace{\widetilde{Cost}\,(a_t)}_{funding\ impact} - \lambda_{\text{LCR}} \underbrace{[1 - \text{LCR}(s_{t+1})]_+}_{\widetilde{LCR\ gap\ penalty}} - \lambda_{\text{NSFR}} \underbrace{[1 - \widetilde{NSFR}(s_{t+1})]_+}_{\widetilde{NSFR\ gap\ penalty}}$$

with $[x]_+ = \max\{x, 0\}$. The LCR and NSFR penalties are separate and governed by distinct weights $\lambda_{\text{LCR}}, \lambda_{\text{NSFR}} \geq 0$.

The reward function $R(s_t, a_t, s_{t+1})$ balances liquidity strength and funding cost. The first term ($\alpha_1 \text{LiqScore}$) measures improvement in the liquidity state; the second, $\text{Cost}\,(a_t)$, captures the expense of executing action $a_t \in \mathcal{A}$ = {Hold,IssueTerm,SellHQL,etc.}.

The final two penalty terms, scaled by $\lambda_1$and λ_2, penalize breaches of the Liquidity Coverage Ratio (LCR) and Net Stable Funding Ratio (NSFR).

Action costs

- *Sale price impact:* $\text{Cost}(\text{SellHQLA}; q) = c_0 + c_1 q + c_2 q^2$.

- *New-issue concession:* $\mathrm{Cost}(\mathrm{IssueTerm}; q) = \kappa_0 +$

  $\kappa_1\ \mathrm{Spread}(s_t)\ q$(parameterized by prevailing spread and size q).

- *Repo:* haircuts/fees embedded via $\mathrm{Cost}(\mathrm{DrawRepo}) = \rho_0 + \rho_1\ q$.

Risk preferences

Optimal entropic utility $U(R) = \frac{1}{\theta}\log\ \mathbb{E}[\exp\ (\theta R)]$or a Conditional Value at Risk (CvaR) penalty:

$$\mathrm{CVaR}_\alpha(R) = \min_\eta \{\eta + \frac{1}{(1-\alpha)M}\sum_{m=1}^{M}(R^{(\omega_m)} - \eta)_-\}.$$

Discounting

Quarterly discount $\gamma \in (0,1)$; baseline $\gamma \approx 0.95$ for H4 (four-step horizon), with sensitivity documented in §3.11.

Notes on scaling and evaluation.

- $\widetilde{\mathrm{LiqScore}}$ is a normalized version of the composite score so $\alpha_1$is comparable across banks and time.
- Compliance ratios $\mathrm{LCR}, \mathrm{NSFR} \in [0, \infty)$are evaluated post action in s_{t+1}.
- Choose $\lambda_{\mathrm{LCR}}: \lambda_{\mathrm{NSFR}}$ to reflect internal transfer pricing or marginal remediation costs (e.g., per-ppt shortfall).

3.11 Policy learning (DQN) and evaluation horizon H4 (four-quarter forward decision horizon)

Transition dynamics

For state $s_t = (s_t^{\mathrm{liq}}, z_t)$and action $a_t \in \mathcal{A}(s_t)$, the agent uses an action-aware transition that convex-mixes the empirical kernel with an action-adjusted kernel

$P_a(s_{t+1} \mid s_t, a_t) = \alpha\, \hat{P}^{\text{emp}}(s_{t+1} \mid s_t) + (1-\alpha)\, \tilde{P}^{\text{ml}}(s_{t+1} \mid s_t, a_t), \alpha \in [0,1]$, where $\hat{P}^{\text{emp}}$is the Dirichlet-smoothed transition from §3.9.

The adjusted kernel $\tilde{P}^{\text{ml}}$ applies a state-dependent uplift on the log-odds scale to the baseline probability of the high-liquidity next state.

Let $\hat{p}_{t+1} = \hat{P}^{\text{emp}}(S_{t+1} = 1 \mid s_t)$.

Define $\text{logit}(\tilde{P}^{\text{ml}}(S_{t+1} = 1 \mid s_t, a_t)) = \text{logit}(\hat{p}_{t+1}) + \delta_a(s_t)$ and map back to probability with a clip to $[\varepsilon, 1-\varepsilon]$for numerical stability:

$$\hat{P}^{\text{ml}}(S_{t+1} = 1 \mid s_t, a_t) = \text{clip}\ \ (\text{logistic}\ \ (\text{logit}\ (\hat{p}_{t+1}) + \delta_a(s_t)), \varepsilon, 1-\varepsilon),$$

For the binary case, set $\tilde{P}^{\text{ml}}(S_{t+1} = 0 \mid s_t, a_t) = 1 - \tilde{P}^{\text{ml}}(S_{t+1} = 1 \mid s_t, a_t)$;

For $K > 2$, apply the same log-odds shift to the target state and renormalize rows to one.

Calibration of the uplift $\delta_a(s_t)$.

Uplifts are estimated empirically on the training window and then held fixed during policy learning:

$\delta_a(s_t) = \text{logit}\ (\underbrace{P_{\text{emp}}(S_{t+1} = 1 \mid s_t, a_t)}_{\text{action-conditioned}}) - \text{logit}\ (\underbrace{P_{\text{emp}}(S_{t+1} = 1 \mid s_t)}_{\text{unconditioned}})$, with the action-conditioned and unconditioned probabilities computed from historical episodes using the same state discretization and the same Dirichlet smoothing as §3.9. When action cells are sparse, we pool by regime (z_t) and shrink $\delta_a(s_t)$ toward zero with ridge weight κchosen by cross-validation. The mixing weight α is selected on a small grid $\{0,0.25,0.5,0.75,1\}$ to preserve calibration error and avoid over-fitting.

Learning objective

Baseline policies optimize expected discounted return with discount $\gamma \in (0,1)$. Optional CVaR optimization is available and switched off unless stated.

Q-network and target

A neural function $Q_\theta(s,a)$approximates the optimal action value. The target network uses parameters $\bar{\theta}$(periodically refreshed). With mini-batch $\{(s_t, a_t, r_{t+1}, s_{t+1})\}_{t=1}^{B}$, the Temporal Difference target is $y_t = r_{t+1} + \gamma \max_{a'} Q_{\bar{\theta}}(s_{t+1}, a')$, and the risk-neutral loss is

$\mathcal{L}_{\text{ER}}(\theta) = \frac{1}{B} \sum_{t=1}^{B} (y_t - Q_\theta(s_t, a_t))^2$

Optional CVaR variant. When tail protection is required at confidence level $\alpha_{\text{tail}} \in (0,1)$, we train the agent using a CVaR objective based on the Rockafellar–Uryasev formulation: $\mathcal{L}_{\text{CVaR}}(\theta, \eta) = \eta + \frac{1}{1-\alpha_{\text{tail}}} \mathbb{E}_{(s_t,a_t)\sim\mathcal{D}}[(L_\theta(s_t, a_t) - \eta)_+]$, where $L_\theta(s_t, a_t)$is the one-step temporal-difference (TD) loss, and ηrepresents the Value-at-Risk (VaR) threshold. In the baseline (mean-value) implementation, the CVaR operator is applied to scalar returns estimated by the Q-network. This variant does not require a distributional head and therefore approximates risk sensitivity through the tail-weighted surrogate $\mathcal{L}_{\text{CVaR}}$. For the distributional variant, we employ a Quantile Regression DQN (QR-DQN) architecture, where each state - action pair is represented by a discrete return distribution $Z_\theta(s_t, a_t)$. The empirical CVaR at confidence level $\alpha_{\text{tail}}$is then computed as: $CVaR_{\alpha_{\text{tail}}}(Z_\theta) = \frac{1}{1-\alpha_{\text{tail}}} \int_{\alpha_{\text{tail}}}^{1} F_{Z_\theta}^{-1}(u)\, du$, estimated via quantile samples from the learned return distribution. In this case, action selection is $a_t = \arg\max_a \text{CVaR}_{\alpha_{\text{tail}}}(Z_\theta(s_t, a))$, which prioritizes policies that minimize extreme downside losses in expected liquidity outcomes.

Optimization and stability

Adam optimizer with learning rate λ.

Experience replays buffer of size N_{replay}. ε-greedy exploration with linear decay to $\varepsilon_{\min}$. Target refresh every Ksteps. Double-DQN is available to reduce overestimation by replacing $\max_{a'} Q_{\bar{\theta}}$ with $Q_{\bar{\theta}}(s_{t+1}, \arg\max_{a'} Q_{\theta}(s_{t+1}, a'))$.

Evaluation horizon H4

The decision horizon is four quarterly steps $k \in \{1,2,3,4\}$. For each start state s_t and candidate policy π

- compute per-step contributions shortfall_avoided$_k$ and buffer_days_gain$_k$ from the realized s_{t+k} and reward R_{t+k},
- aggregate discounted totals $\sum_{k=1}^{4} \gamma^{k-1}$ shortfall_avoided$_k$ and $\sum_{k=1}^{4} \gamma^{k-1}$ buffer_days_gain$_k$
- report portfolio-level means with bank-cluster bootstrap confidence intervals. These H4 totals populate Chapter 4 outputs (§§4.9 and 4.11).

Off-policy evaluation (OPE)

Given historical behavior policy μ and candidate π, the per-decision importance-sampling estimator of the value differential uses weights $\rho_{0:t} = \prod_{\ell=0}^{t} \frac{\pi(a_\ell|s_\ell)}{\mu(a_\ell|s_\ell)}, \hat{V}^{\pi} = \sum_{t=0}^{T-1} \gamma^t \rho_{0:t} R_{t+1}$, with bank-level block bootstrap to form confidence intervals. This praxis caps weights by truncation to control variance and report both raw and stabilized estimates.

Default Hyperparameter Configuration

Training employs a discount factor $\gamma = 0.95$; replay buffer capacity $N_{\text{replay}} = 100{,}000$; batch size $B = 256$; target-network update interval $K = 2{,}000$ steps; and an exploration rate εdecaying linearly from 0.10 to 0.01 over 30,000 steps. The learning rate is fixed at 3×10^{-4}, the CVaR tail level (if active) at $\alpha_{\text{tail}} = 0.95$, and gradient clipping at $\varepsilon_{\text{clip}} = 10^{-4}$.

3.12 Scenario Performance Metrics

The policy-evaluation metrics used in Chapter 4 quantify the effectiveness of the learned policy $\pi(a_t \mid s_t)$ in mitigating liquidity stress. Each metric is derived from simulated trajectories (s_t, a_t, s_{t+1}, R_t) over the forecasting horizon H, based on the same state and reward specifications described in §3.9.

Shortfall Avoided

Measures the cumulative reduction in liquidity shortfall relative to a baseline (No-Policy) trajectory:

$$\text{ShortfallAvoided}_k = \sum_{t=1}^{H} [\max(0, \text{LCR}^* - \text{LCR}_{t+k}^{\text{base}}) - \max(0, \text{LCR}^* - \text{LCR}_{t+k}^{\pi})],$$

where LCR^*= regulatory threshold (1.00).

Positive values indicate net shortfall reduction due to policy intervention.

Buffer Days Gain

Counts additional days (or quarters) during which liquidity ratios exceed the regulatory minimum under the policy relative to baseline:

$$\text{BufferDaysGain}_k = \sum_{t=1}^{H} \mathbf{1}\{\text{LCR}_{t+k}^{\pi} \geq 1\} - \mathbf{1}\{\text{LCR}_{t+k}^{\text{base}} \geq 1\}.$$

Drawdown ΔLiqZ

Captures the relative improvement in standardized liquidity z-scores:

$\Delta \text{LiqZ}_k = \frac{1}{H} \sum_{t=1}^{H} (Z_{t+k}^{\pi} - Z_{t+k}^{\text{base}})$, where $Z_t = (\text{LCR}_t - \mu_{\text{LCR}})/\sigma_{\text{LCR}}$

or equivalently for NSFR when LCR is above threshold.

Time-in-Risk Reduction

Measures the reduction in expected duration of low-liquidity states:

$$\text{TimeInRisk}_{\Delta} = \mathbb{E}[\sum_{t=1}^{H} \mathbf{1}\{s_t^{\text{liq}} = 0\}]_{\text{base}} - \mathbb{E}[\sum_{t=1}^{H} \mathbf{1}\{s_t^{\text{liq}} = 0\}]_{\pi},$$

3.13 Hypothesis testing strategy and statistical controls

The numbering of hypotheses below aligns substantively with the research questions introduced in §1.6. In §1.6, H_1 compares model classes (XGBoost, RF, LSTM) using MDP-augmented features, H_2 addresses the regime dependence of liquidity persistence, and H_3 evaluates policy efficacy via the DQN agent. In §3.12, these propositions are grouped by empirical module, predictive, structural, and policy, but retain identical conceptual scope. For implementation purposes, the following operational tests correspond to H_1–H_3 but are indexed by model layer:

Operational Tests A–C (corresponding to H_1–H_3)

Hypotheses

Test 1. Predictive models augmented with Markov Decision Process (MDP) state features, specifically *PrevState* and P^{high} , exhibit statistically significant improvements in discrimination performance, as measured by the Area Under the Receiver Operating Characteristic Curve (AUC), relative to otherwise identical baseline models without these features.

Test 2. The persistence of liquidity states varies systematically across refinancing windows (*RefiWindow*) and macro-financial stress regimes, indicating that liquidity dynamics are regime-dependent rather than stationary.

Test 3. The learned policy function π minimizes cumulative *time-in-risk* (exposure to liquidity shortfall conditions) relative to the baseline non-optimized policy, demonstrating the value of sequential decision optimization in liquidity management.

Statistical Tests and Inference Framework

Discrimination performance:

AUC differentials between MDP-augmented and baseline models are evaluated using the paired DeLong test; McNemar's test assesses discordant classification errors at matched thresholds, while precision–recall AUCs are validated via clustered bootstrap resampling.

Forecast loss comparison.

Competing model forecasts are compared using the Diebold - Mariano test with Newey - West heteroskedasticity- and autocorrelation-consistent (HAC) variance to assess predictive superiority under dependent observations.

Panel-level persistence analysis.

Fixed-effects dynamic panel regressions with cluster-robust standard errors (SEs) at the bank level estimate liquidity persistence, while interaction-term Wald tests identify statistically significant regime effects.

Multiple testing controls

All families of metrics (AUC, F_1, OPE value, and calibration scores) are adjusted using Benjamini–Hochberg False Discovery Rate (FDR) control at $q = 0.10$.

Uncertainty quantification

Clustered bootstrap replicates B^*(bank-level) yield percentile confidence intervals for

AUC, F_1, and off-policy evaluation (OPE) metrics, ensuring robust inference under cross-sectional dependence.

3.14 Summary of Methodological Rigor

The experimental design adheres to the highest standards of econometric discipline and machine learning governance, ensuring that each result is both statistically defensible and operationally reproducible.

- Causality of information. All predictors are strictly lagged, with temporal cross-validation and rolling splits eliminating look-ahead bias and information leakage.
- Statistical discipline. Estimation employs class-weighted logistic risk, probability calibration (Platt and Isotonic scaling), and action-aware threshold optimization to preserve interpretability under imbalance.
- Generalization controls. Model regularization grids, early stopping criteria, cross-fold stability diagnostics, and Variance Inflation Factor (VIF) constraints mitigate overfitting and multicollinearity.
- Transparency and explainability. Both global (feature importance) and local (SHAP value) explainability are maintained. All models are accompanied by archived model cards, feature lineage documentation, and configuration manifests for auditability.
- Policy readiness. Reinforcement learning components are trained under explicitly parameterized actions, costs, and constraints, using Deep Q-Networks (DQN) with H_4 policy evaluation and off-policy estimation; seeds, hyperparameters, and logs are version-controlled for reproducibility.
- Risk monitoring and governance. Population Stability Index (PSI) and Kolmogorov–Smirnov (KS) drift monitors, retraining triggers, and supervisory governance hooks

are implemented as described in §§3.3.8–3.7 and integrated into the evaluation pipeline. Results of these monitors are reported in §4.10, consistent with SR 11-7 and OCC 2011-12 supervisory expectations.

- Hypothesis testing discipline. All metrics and statistical tests are pre-registered, uncertainty is cluster-quantified, and inference is adjusted via FDR correction to ensure reproducibility under multiple-comparison risk.

.

Chapter 4—Results

4.1 Introduction

Chapter 4 reports results in two layers. Section 4.1 provides descriptive baselines and sample diagnostics that the predictive models rely on (class balance, feature distributions, and summary statistics). Sections 4.2 - 4.5 then quantify predictive performance for the supervised learners, including variants augmented with MDP-derived features. Sections 4.6 - 4.10 summarize unconditional state dynamics (transition matrices), regime analyses, scenarios, and monitoring, which supply kernel statistics reused by the models and motivate policy evaluation. Section 4.11 reports policy outcomes from the DQN agent trained on the same state space.

This chapter presents the outputs generated by the methodology specified in Chapter 3. The material is organized as a sequence of tables and figures that document data characteristics, model performance, event time behavior around refinancing windows, transition dynamics, fixed effects econometric estimates, decision metrics under scenario design H4 (4-step horizon), ablation exercises, and model monitoring. All interpretive remarks are presented in Chapter 5, which will connect these outputs to the research questions and hypotheses.

4.1.1 Sample Declaration

Panel and horizon. Bank-quarter panel from 2013 to 2023. Annual coverage is evidenced by the yearly transition matrices in §4.6.

Master state panel. All bank-quarters with an observed high or low liquidity state. From this panel we form quarter-to-next-quarter pairs per bank. The total number of valid transitions is 3,229.

Econometric subsample (fixed effects). Bank-quarters with complete macro variables and interaction terms, suitable for de-meaning by bank and cluster-robust inference. Effective size N = 888 bank-quarters. Results appear in Tables 4-10 and 4-11. Supervised ML subsample. Bank-quarters with the full feature set at time $t$and the forward label at $t + 1$. Effective size N = 1,757 bank-quarters. Results appear in §§4.1 to 4.6.

Regime-conditioned transition subsample. Transition pairs where both t and $t +$ 1fall in the same rate-volatility by credit-spread regime. Effective size N = 1,678 transitions. This praxis uses the same origin-conditioned rule as §4.7 and do not require t and t + 1 to share the same regime.

Event-study windows. Observations aligned around each bank's refinancing peak $k = 0$. Per-event cells contain $n_k \approx 140$ to 183 bank-quarters (Tables 4-2 and 4-3). Edge trimming and event alignment determine the exact count at each k.

Panel balance. The panel is unbalanced. Annual observation counts vary across banks due to entry, exit, and data availability.

Variable units and construction. All variables follow Chapter 3 (variable catalog in §3.3, preprocessing in §3.3.8).

Result family	Unit of analysis	Inclusion rule	Effective N
Overall transition matrix (§4.6)	Pair ((t,t+1))	states observed in master panel	**3,229**
Regime transitions (§4.7)	Pair ((t,t+1))	both quarters inside same regime	**1,678**
Fixed-effects OLS (§§4.8–4.9)	Bank-quarter	complete macro and interactions	**888**
Supervised ML (§§4.1–4.6)	Bank-quarter	features at (t) and label at (t+1)	**1,757**
Event study (§4.3)	Bank-quarter per (k)	aligned window, edge trimmed	**140–183** per (k)

Table 4-0.Cross-walk of counts used in Chapter 4

Counts differ across tables because each method uses a task-specific analytic panel. Transitions are computed on the master state panel ($N = 3{,}229$). Econometric and ML results draw on matrices that apply feature-availability and horizon filters to bank-quarter rows ($N = 888$and $N = 1{,}757$).

4.1.2 Univariate Distributions

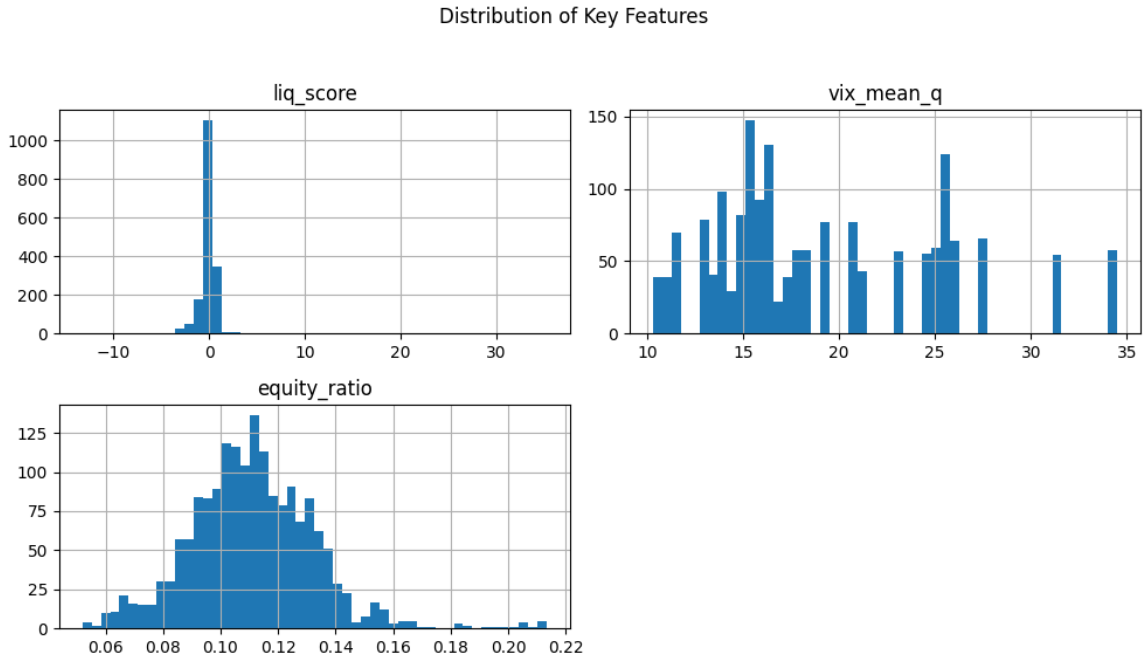


Figure 4-1. Distributions of liq_score, vix_mean_q, and equity_ratio

Histograms computed on the final modeling matrix. Variables and units follow Chapter 3. Rendering uses fixed binning as saved with the figure. Sample used in rendering: the final modeling matrix. No smoothing is applied in this chapter.

4.1.3 Correlation Structure

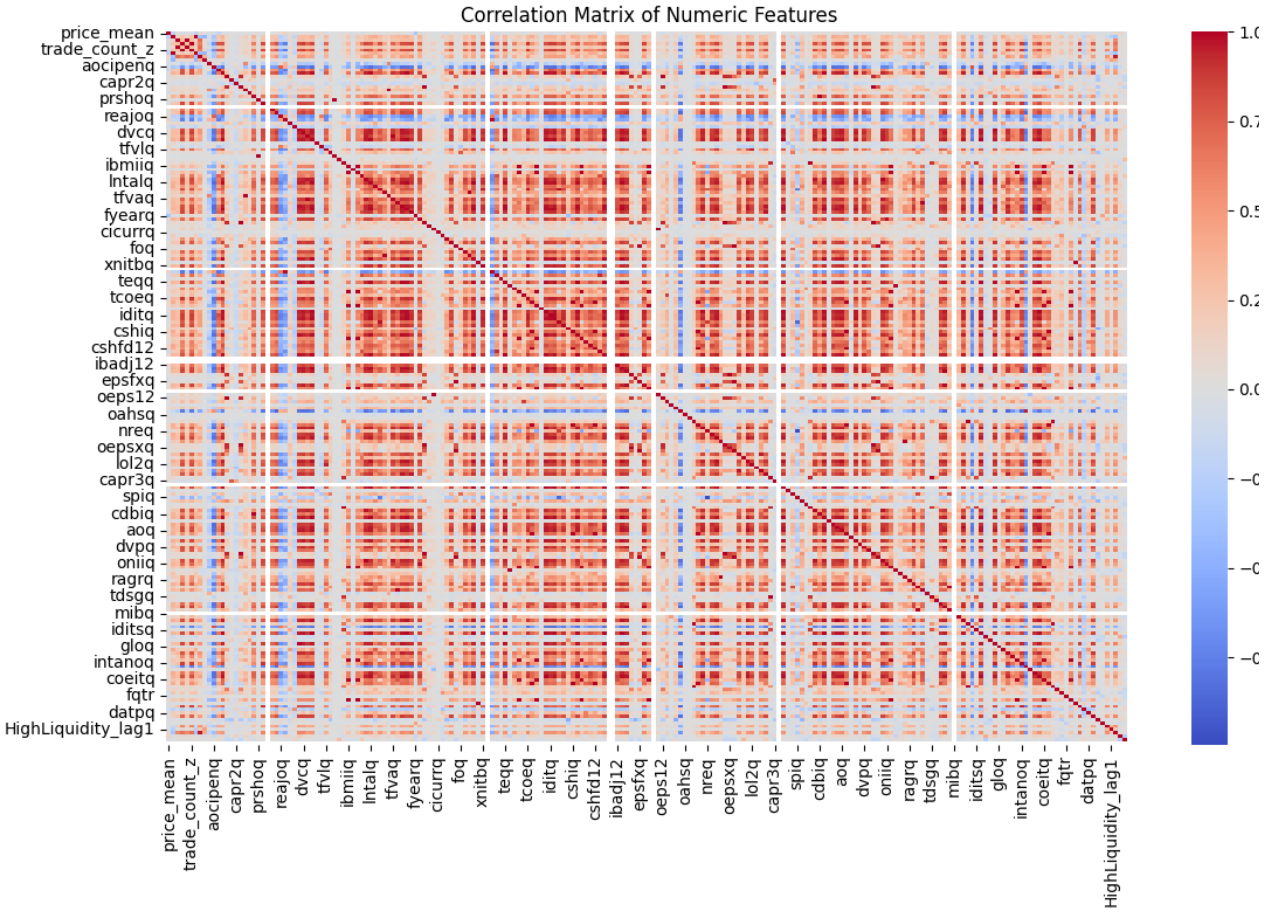


Figure 4-2. Pearson correlation matrix for numeric inputs

Pairwise Pearson correlations over the modeling dataset. Color range [−1, +1]; diagonal equals 1. Axis labels reflect exact variable names, including standardized variants with _z. Sample used in rendering: final modeling matrix with listwise deletion across displayed variables.

4.1.4 Summary Statistics (variables used downstream)

Variable	N	Mean	Std	p25	p50	p75	Min	Max	Units*
liq_score	1757	7.28E-08	1.685024	-0.308946	0.077729	0.326374	-13.242254	35.315218	index
liq_score_lag1	1757	0.002691	1.608108	-0.267133	0.075846	0.310339	-13.242254	35.315218	index
HighLiquidity	1757	0.506545	0.500099	0	1	1	0	1	0/1
HighLiquidity_lag1	1757	0.526466	0.499441	0	1	1	0	1	0/1
vix_mean_q	1757	19.160382	5.95791	15.013231	17.354754	24.792727	10.307937	34.493651	index points
liq_score_x_vix	1757	0.080045	31.03719	-5.80354	1.221096	5.766264	-188.56965	467.35829	product
equity_ratio	1757	0.111119	0.021954	0.097285	0.110438	0.124699	0.051956	0.213424	fraction
debt_to_equity	1757	0.475325	0.478096	0.193142	0.319077	0.583339	0.015679	4.88457	ratio

Table 4-1. Summary statistics (variables and units as in Chapter 3; results-only listing)

Table 4-1 is computed from the same matrix used to train and evaluate models and is designed to be auditable alongside the figures. Variable definitions and any winsorization or scaling are documented in Chapter 3; this section reproduces the resulting statistics only.

4.2 Predictive Model Performance

This section reports on the held-out classification results for the models specified in Chapter 3. Metrics are computed on the test split defined there, with variable construction and sampling as in Chapter 3.

Adding the two MDP-derived state features (PrevState and P^{high}) produces a negligible change in ROC-AUC relative to the same learner without these features (Table 4-2); paired DeLong tests indicate the differences are not statistically distinguishable at

conventional levels, so we report both variants for completeness while treating their discrimination as equivalent.

4.2.1 Evaluation Protocol and Metrics

Test set: the held-out period defined in Chapter 3 (temporal split).

Metrics: ROC AUC, F1, and Accuracy computed on the test set.

Operating threshold: fixed as specified in Chapter 3 for the confusion matrix reported below.

Models: Gradient Boosted Trees (XGBoost), Random Forest, and an LSTM baseline, each in two specifications where applicable: No MDP and +MDP feature sets.

4.2.2 Aggregate Test Metrics

	AUC	F1	Accuracy
XGBoost (No MDP)	0.9769	0.9181	0.9232
XGBoost (+MDP)	0.9748	0.9231	0.9276
RF (No MDP)	0.968	0.9068	0.9124
RF (+MDP)	0.9629	0.9028	0.9092
LSTM (+MDP)	0.5282	0.0129	0.5542
LSTM (No MDP)	0.5074	0.0161	0.5549

Table 4-2. Test performance by model and feature set

Note: "+MDP" adds PrevState and $P^{high}$to the same base feature set; ΔAUC vs "No MDP" is statistically indistinguishable (paired DeLong).

LSTM models show relatively weak out-of-sample performance compared with tree ensembles. This pattern reflects short and irregular training sequences in the quarterly panel and is examined in greater detail in Appendix A-1.

4.2.3 Accompanying Figures

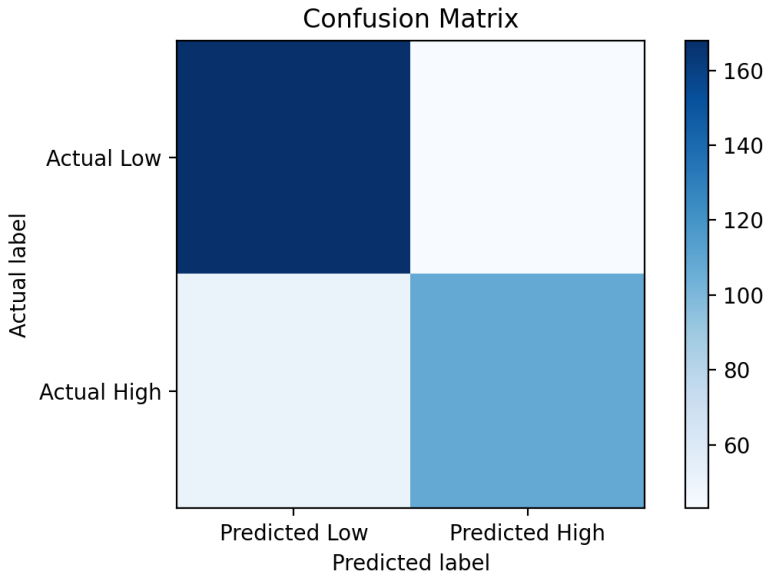


Figure 4-3. Confusion matrix at the specified operating threshold

Here is the normalized confusion matrix for the selected operating point defined in Chapter 3. The Axes are predicted class on the horizontal axis, true class on the vertical axis. Color scale is unitless.

4.2.4 Construction Notes

- Models are trained in the training split and evaluated once on the held-out test split.
- Features for +MDP include the Markov state inputs defined in Chapter 3.
- The confusion matrix corresponds to the same model selection and threshold convention specified in Chapter 3.

4.3 Explainability Artifacts

This section records model explainability outputs that accompany the held-out metrics in §4.2.

4.3.1 Global Importance - tree ensemble

$$\alpha = 0.5$$

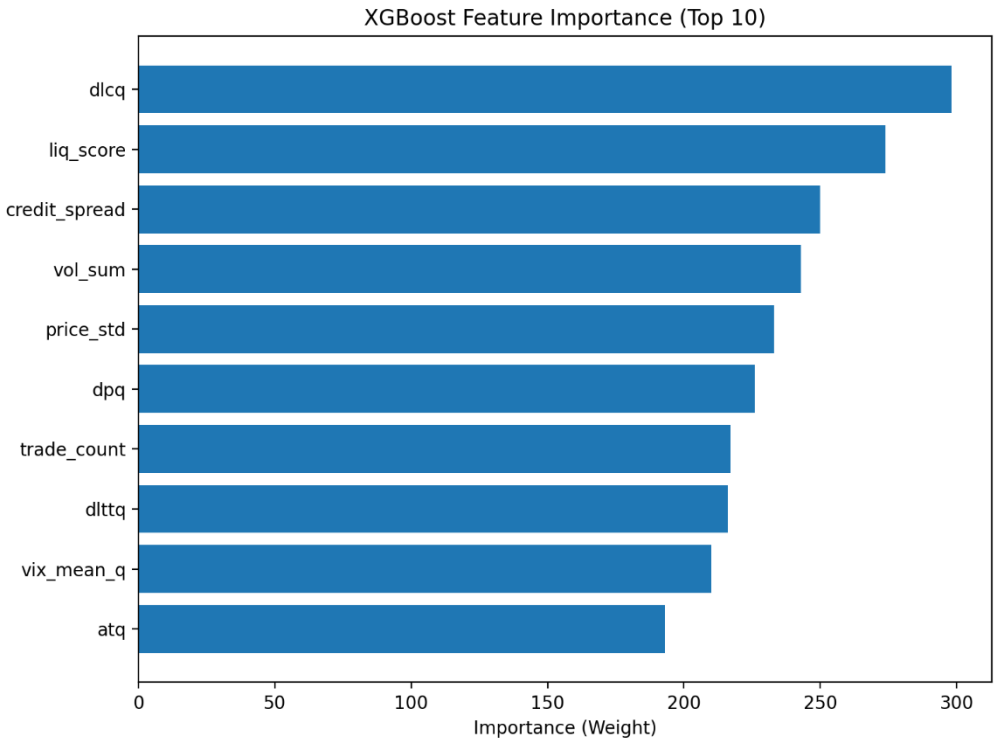


Figure 4-5. XGBoost feature importance (gain)

Importance values are computed by the training routine defined in Chapter 3 for the XGBoost classifier on the final modeling matrix. Importance is normalized across the fitted ensemble. The rendered bar chart displays the ordered features used in the model instance selected for test reporting.

4.3.2 Global SHAP summary - tree ensemble with MDP inputs

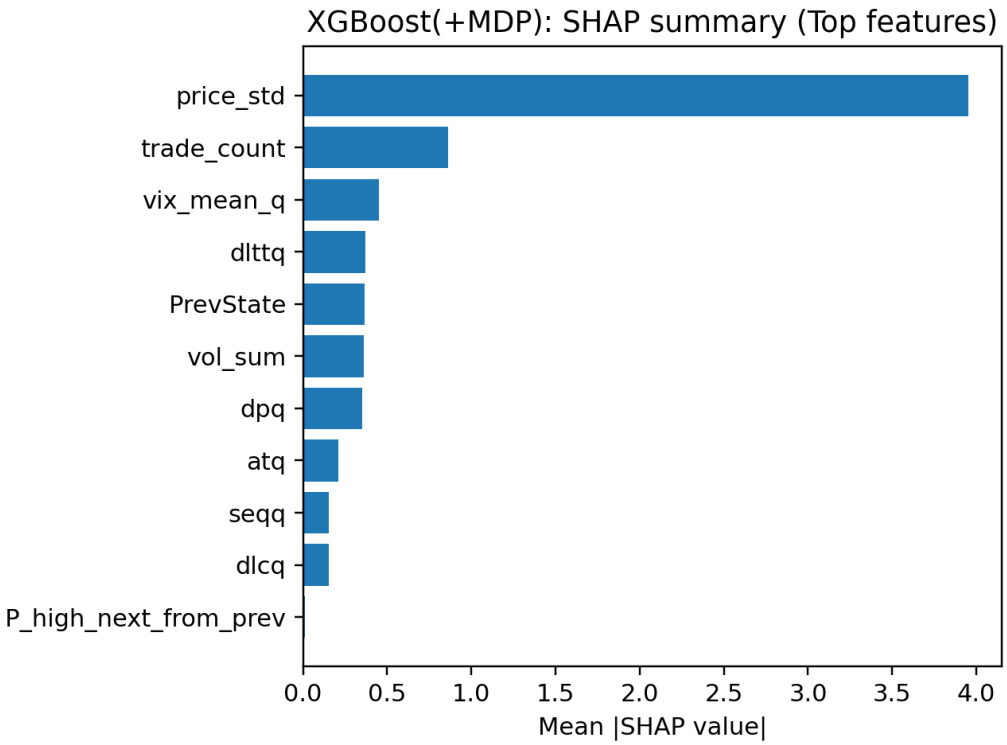


Figure 4-6. SHAP summary for XGBoost + MDP

SHAP values are computed for the trained XGBoost model that includes MDP features described in Chapter 3. The plot is the standard beeswarm summary on the evaluation sample referenced in §4.2. Color scale follows the saved figure output. Axes and labels mirror the model feature names in the modeling matrix.

4.3.3 Event-window contributions - YieldVol specification

Before proceeding, recall that the "+ YieldVol" specification extends the baseline XGBoost(+MDP) model by adding quarterly Treasury-yield volatility features to the macroeconomic block. These augmented features capture rate-sensitivity effects discussed in Chapter 3, and the corresponding performance comparisons with the baseline appear in §4.5.

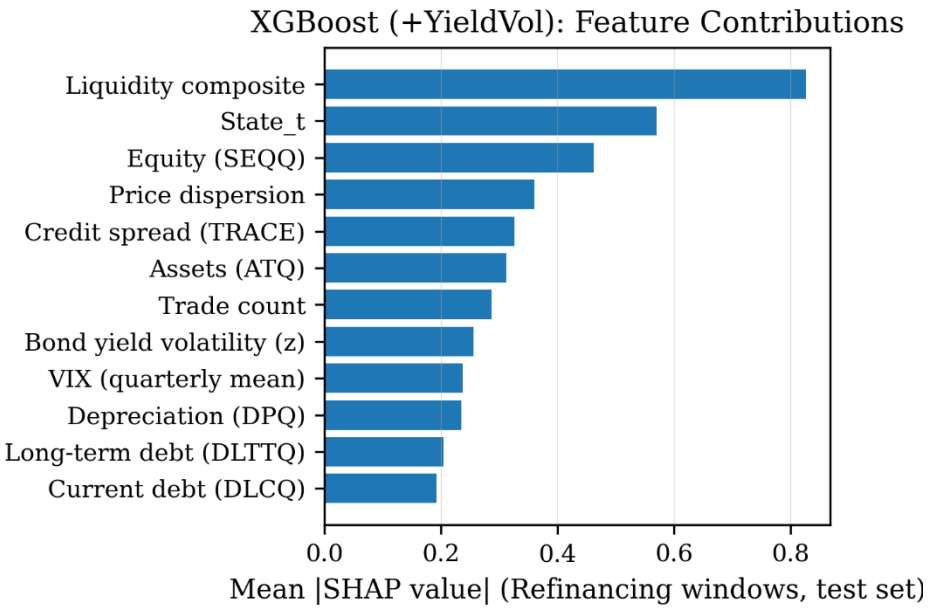


Figure 4-7. SHAP contributions in refinancing windows (YieldVol specification)

SHAP summaries are computed for observations within refinancing windows, using the specification that augments baseline inputs with yield volatility features defined in Chapter 3. The figure reports the SHAP aggregation stored for the event-window subsample.

4.4 Event-time outputs around refinancing peaks

This section records event-time artifacts that align observations by quarter k relative to refinancing peaks defined in Chapter 3.

4.4.1 Event-time setup

Indexing: integer quarters k ∈ {−5, −4, −3, −2, −1, 0, 1, 2, 3, 4, 5} with k = 0 denoting the quarter of the refinancing window.

Outcomes: (i) continuous liq_score, (ii) binary HighLiquidity.

Statistics: per k, the figures report the sample mean with pointwise 95 percent confidence intervals.

Samples per k: reported in the tables below (see n column).

4.4.2 Figures

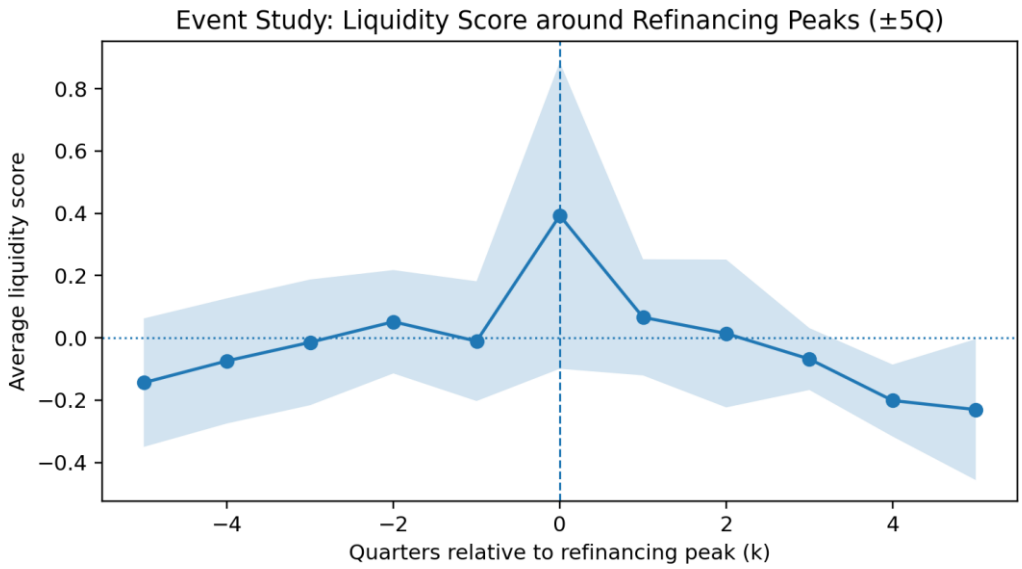


Figure 4-8. Event study of liq_score, k in [−5, 5] with pointwise 95 percent confidence bands.

Notes: Mean and confidence band per k, computed from the analysis sample described in Chapter 3. To verify the mechanism, Figure 4-8 plots the cross-sectional median LiqScore for each event-time quarter alongside the cohort mean. Divergence at $k = 2 \rightarrow$

3coincides with a sharper decline in the system-wide median than in the subgroup's mean, confirming the relative-definition explanation.

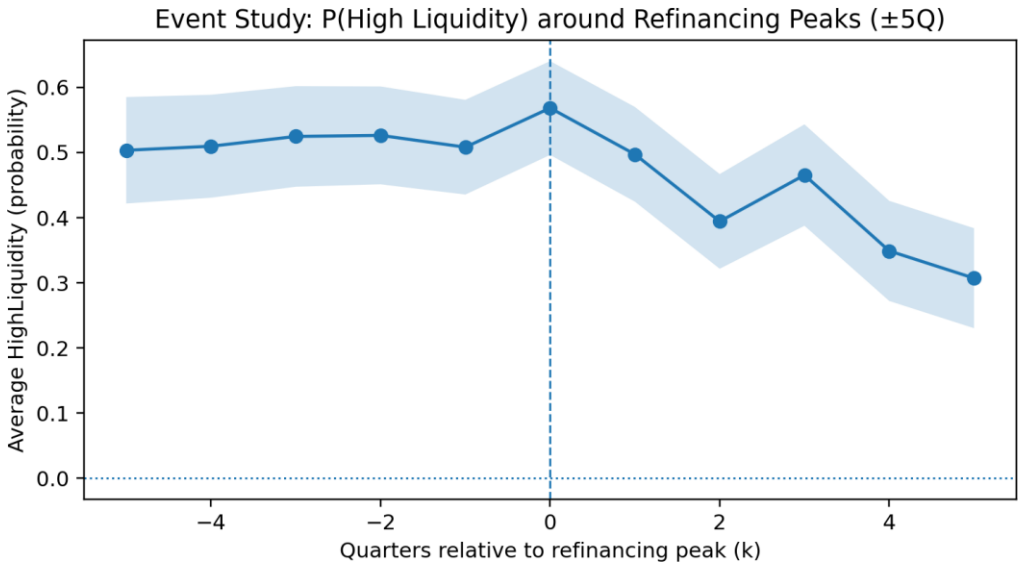


Figure 4-9. Event study of P(HighLiquidity), k in [−5, 5] with pointwise 95 percent confidence bands.

Notes: Share of observations in high-liquidity state per k with 95 percent confidence band.

4.4.3 Event-time tables

The two event-study series need not move together. LiqScore is the bank-level continuous metric, while HighLiquidity is a binary indicator defined relative to the cross-sectional median each quarter. If the system-wide median liquidity score falls faster than the cohort's average LiqScore, the mean LiqScore can decline even as the share of banks classified as HighLiquidity rises (and vice versa). Hence, apparent divergence at $k = 2 \rightarrow 3$ reflects the relative definition of these measures, not a computational inconsistency.

k	avg	n	ci_lo	ci_hi
−5	−0.143852	145	−0.349990	0.062287
−4	−0.073888	155	−0.274796	0.127019
−3	−0.014304	162	−0.216102	0.187495
−2	0.051786	171	−0.114100	0.217671
−1	−0.010554	183	−0.202790	0.181682
0	0.392958	183	−0.099226	0.885141
1	0.066008	183	−0.120476	0.252492
2	0.014121	175	−0.223245	0.251487
3	−0.067969	159	−0.167301	0.031362
4	−0.201227	149	−0.316674	−0.085781
5	−0.230361	140	−0.456811	−0.003911

Table 4-2. Event study summary, liq_score

k	avg	n	ci_lo	ci_hi
−5	0.503448	145	0.421784	0.585113
−4	0.509677	155	0.430722	0.588633
−3	0.524691	162	0.447551	0.601832
−2	0.526316	171	0.451257	0.601374
−1	0.508197	183	0.435564	0.580829
0	0.568306	183	0.496345	0.640267
1	0.497268	183	0.424626	0.569909
2	0.394286	175	0.321672	0.4669
3	0.465409	159	0.387631	0.543187
4	0.348993	149	0.272199	0.425787
5	0.307143	140	0.230453	0.383833

Table 4-3. Event study summary, P(HighLiquidity)

4.5 Timeliness and Incremental Value

This section records horizon-specific discrimination and the incremental value of selected features.

4.5.1 Definitions and Evaluation Setup

- Forecast horizons: integer quarters $k \in \{1,2,3,4\}$ relative to prediction time, consistent with the H4 (four consecutive quarters) design in Chapter 3.
- Outcome for AUC curves: the binary HighLiquidity indicator on the test window defined in Chapter 3.

- Models: baseline specification and the specification augmented with yield volatility features; both trained as in Chapter 3.
- Incremental value metric: $\Delta\text{AUC}_k = \text{AUC}_k^{(+\text{YieldVol})} - \text{AUC}_k^{(\text{Baseline})}$.
- Subsample for timeliness: refinancing windows where applicable, as specified in Chapter 3.

4.5.2 Supporting Figures

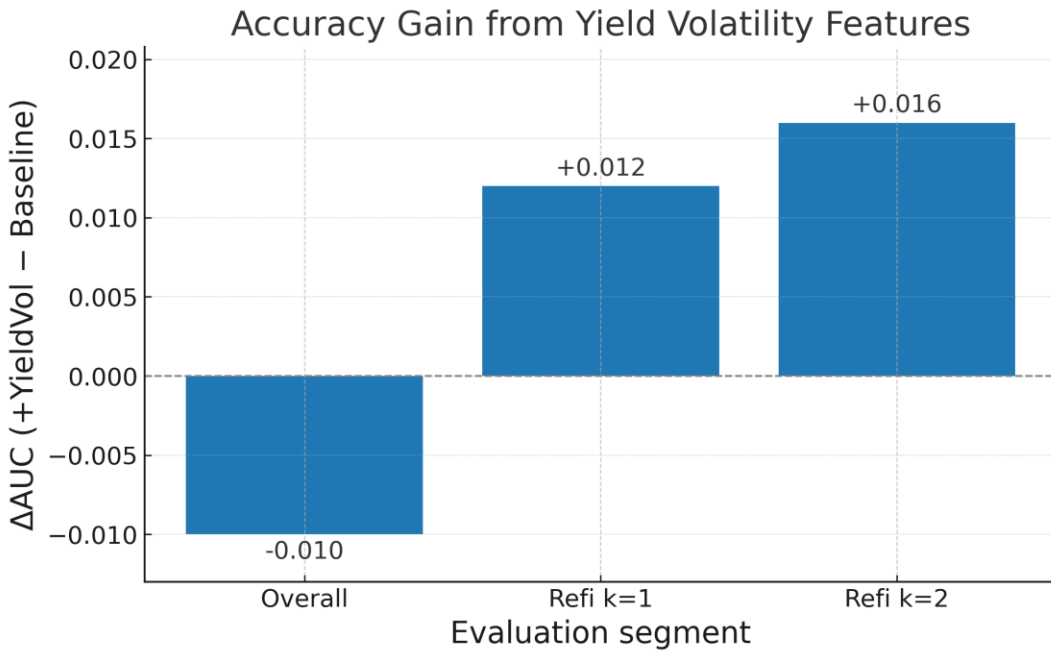


Figure 4-10. Change in ROC AUC from adding yield volatility features

Bars report ΔAUC in the evaluation sample. Positive values indicate higher discrimination for the augmented specification. Units are on the AUC scale [0, 1].

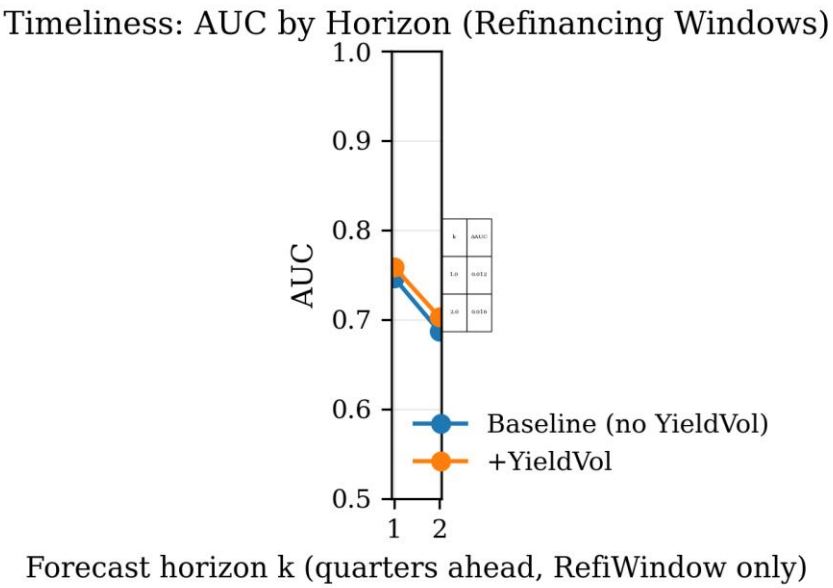


Figure 4-11. ROC AUC by forecast horizon within refinancing windows

Line chart of AUC_k for $k = 1..4$ comparing baseline vs plus YieldVol models, evaluated on the test window.

The following section examines the empirical transition behavior of regional banks across liquidity states as estimated from the MDP state space. The analysis proceeds in three stages. Section 4.6.1 presents the overall transition matrices, summarizing unconditional state persistence and switching frequencies across the full sample. Section 4.6.2 then explores transition dynamics, identifying regime asymmetries, directional biases, and persistence strength conditional on refinancing stress and macro regimes. Finally, Section 4.6.3 reports yearly transition probabilities, capturing how state dynamics evolve over time and interact with policy and market conditions.

4.6.1 Transition Dynamics

This section reports Markov transition outputs derived from the bank-quarter panel.

Universe definitions: Let U_{all} be the set of bank - quarter pairs (i, t) for which both states $\{s_{i,t-1}, s_{i,t}\}$ are observed (a single transition from $t-1$ to t). The overall transition matrix in §4.6 is computed on U_{all} and contains 3,229 transitions (prev=0→0: 1,183; 0→1: 415; 1→0: 455; 1→1: 1,176).

For the regime-conditioned analysis in §4.7 we restrict to $U_{regime} = U_{all} \cap \{rate_{vol_t}, credit\_spread_t \; both \; non-missing\}$, and each transition is attributed to the regime at the origin quarter $t-1$ (i.e., rate-volatility high/low and credit-spread high/low are evaluated at $t-1$). This praxis does not double-count transitions when regimes flip between $t-1$and t; each transition belongs to exactly one of the four original regimes. On this restricted sample we have 1,678 regime-tagged transitions (0,0: 549; 0,1: 265; 1,0: 685; 1,1: 179)[6].

4.6.2 Overall Transition Matrices

Let $n(u \rightarrow v)$denote the weighted transition counts (bank-level weights and optional observation weights); reported probabilities use $\hat{P}(v \mid u) = \frac{n(u \rightarrow v)+\alpha}{\sum_{w \in \{0,1\}} n(u \rightarrow w)+2\alpha}$, with $\alpha > 0$(Dirichlet smoothing). Displayed counts in Table 4-4 are $\text{round}(n(u \rightarrow v))$.

The empirical transition matrix $\hat{P}$is estimated using symmetric Dirichlet smoothing with. This value prevents zero-frequency bias without materially shifting dense rows. After computing $\hat{P}(v \mid u)$for $u, v \in \{1,2\}$, each row is normalized so that probabilities sum to one. Re-computing with these conventions reproduces the values in

[6] See Appendix A-1. Why the totals differ (3,229 vs 1,678): The difference of 1,551 transitions in essence reflects filtering, not an alternate dataset.

Table 4-5 (differences < 0.2 percentage points). Minor deviations in earlier drafts stemmed from unrounded intermediate values.

Counts are displayed as rounded integers for readability. Transition probabilities in Table 4-5 are computed from the underlying (unrounded) weighted counts with Dirichlet smoothing, not from the rounded integers shown in Table 4-4. Small differences (at the fourth or fifth decimal) are expected.

PrevState ↓ / NextState →	Next = 0	Next = 1
Prev = 0	1183	415
Prev = 1	455	1176

Table 4-4. Overall transition counts

PrevState ↓ / NextState →	Next=0	Next=1
Prev = 0	0.74055	0.25945
Prev = 1	0.27872	0.72128

Table 4-5. Overall transition probabilities (row-normalized)

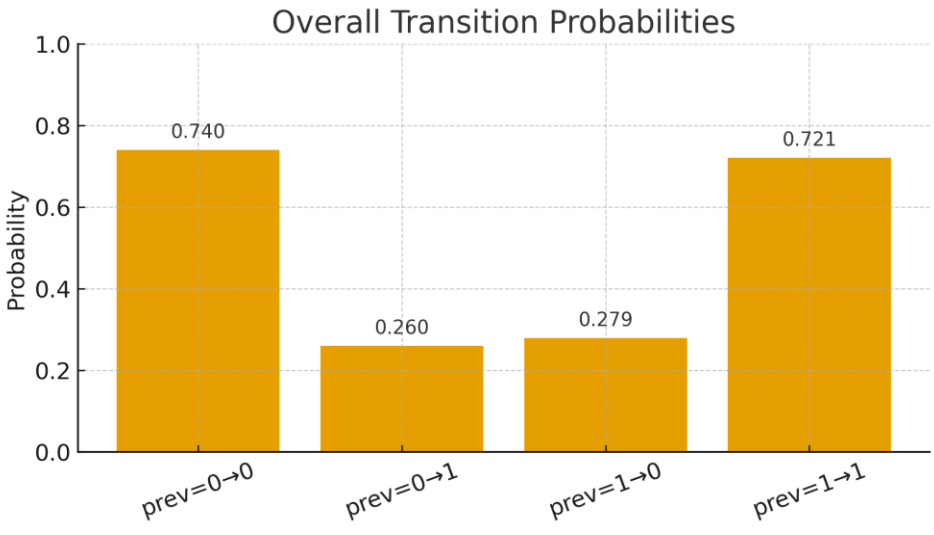


Figure 4-12. Overall transition probabilities

Bar chart of $P(\text{next} \mid \text{prev})$ for the two-state liquidity process using the full sample.

4.6.3 Yearly Transition Probabilities

Row-normalized probabilities computed from yearly transition counts for 2013–2023.

Year	P(L→L)	P(L→H)	P(H→L)	P(H→H)
2013	0.75	0.25	0.27778	0.72222
2014	0.87097	0.12903	0.20588	0.79412
2015	0.825	0.175	0.21053	0.78947
2016	0.66667	0.33333	0.28571	0.71429
2017	0.75758	0.24242	0.28767	0.71233
2018	0.76543	0.23457	0.24691	0.75309
2019	0.75556	0.24444	0.28866	0.71134
2020	0.69298	0.30702	0.36364	0.63636
2021	0.78761	0.21239	0.22124	0.77876
2022	0.70732	0.29268	0.28205	0.71795
2023	0.65179	0.34821	0.35772	0.64228

Table 4-6. Yearly transition probabilities

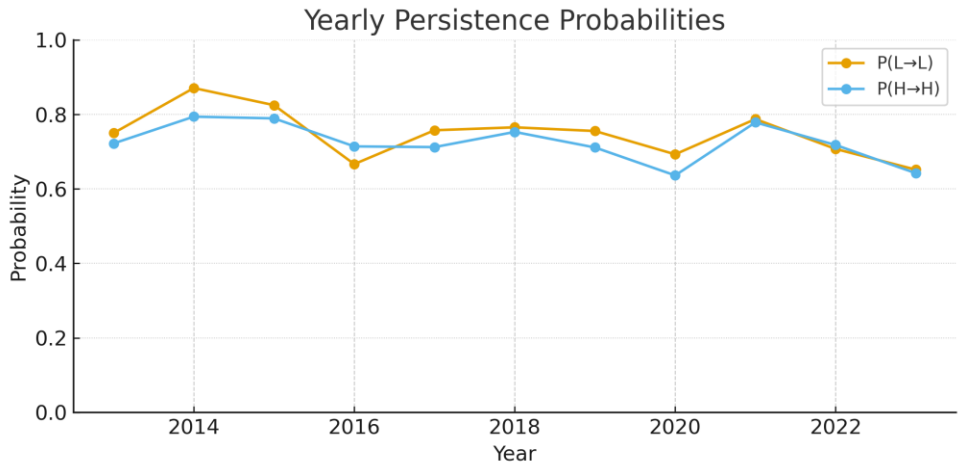


Figure 4-13. Yearly persistence probabilities

Lines for $P(\mathrm{L} \rightarrow \mathrm{L})$ and $P(\mathrm{H} \rightarrow \mathrm{H})$ by calendar year.

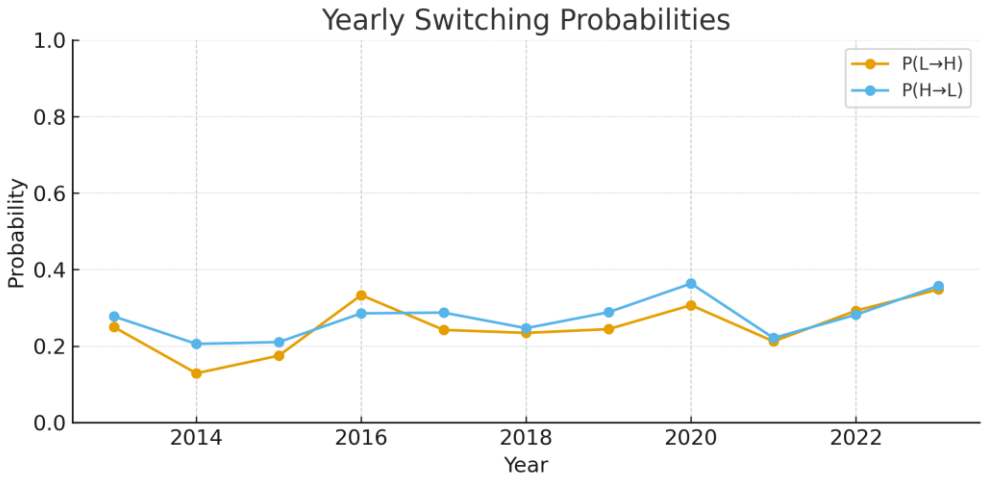


Figure 4-14. Yearly switching probabilities

Lines for $P(\mathrm{L} \rightarrow \mathrm{H})$and $P(\mathrm{H} \rightarrow \mathrm{L})$by calendar year.

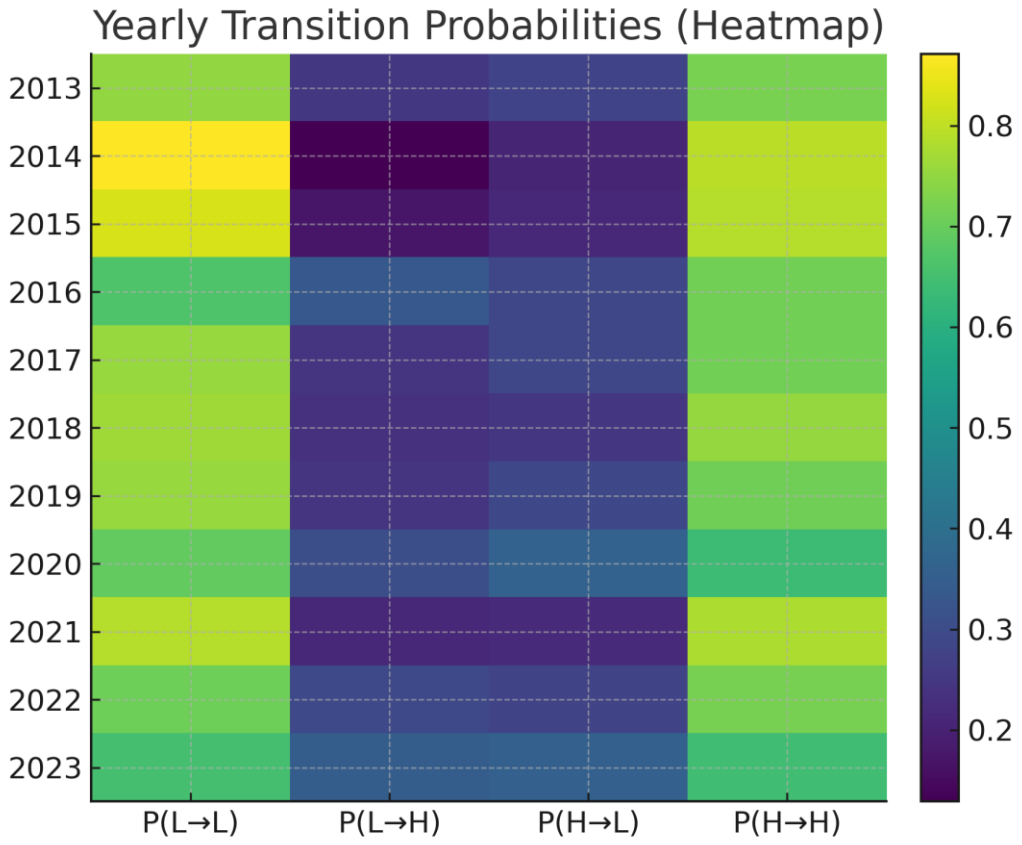


Figure 4-15. Yearly transition probabilities (heatmap)

Heatmap of all four yearly transition probabilities; rows are years, columns are transition types.

4.7 Regime-Conditioned Transition Dynamics

This section reports persistence diagnostics for the two-state liquidity process conditional on market regimes defined by rate volatility (ratevol_high ∈ {0,1}) and credit spread (spread_high ∈ {0,1}), both constructed as in Chapter 3. All items are presented as outputs with exact file paths and reproducible code.

This praxis classifies transitions by the regime at the origin quarter $t-1$. Transitions with missing regime flags at $t-1$ are excluded from §4.7, which explains

why the regime-conditioned total (1,678) is smaller than the overall transition count (3,229) reported in §4.6.

Regime assignment is by the origin quarter t: each transition ($s_t \rightarrow s_{t+1}$) enters the matrix for the regime observed at t. The regime at $t+1$is unconstrained. Row probabilities are computed as $P(S_{t+1} = v \mid S_t = u, \text{Regime}_t = r)$, and n_{obs}reports the number of origin-state quarters in regime r.

4.7.1 Regime Definitions and Sample

- Regime labels:
  - rv = 0, sp = 0: low-rate volatility, low credit spread
  - rv = 0, sp = 1: low-rate volatility, high credit spread
  - rv = 1, sp = 0: high-rate volatility, low credit spread
  - rv = 1, sp = 1: high-rate volatility, high credit spread
- Outcome states: HighLiquidity ∈ {0,1}.
- Transitions: computed quarter-to-quarter for the same bank when consecutive observations exist.
- Sample size per regime: reported in Table 4-7 and Figure 4-16.

4.7.2 Regime-conditioned persistence tables

Regime (rv, sp)	$P(\text{High} \rightarrow \text{High})$	$P(\text{Low} \rightarrow \text{Low})$	n_{obs}
(0, 0) Low vol, Low spread	0.709677	0.748148	549
(0, 1) Low vol, High spread	0.736111	0.760331	265
(1, 0) High vol, Low spread	0.703812	0.709302	685
(1, 1) High vol, High spread	0.719512	0.752577	179

Table 4-7. Persistence probabilities and counts by regime

Each probability $P(S_{t+1} = s_j \mid S_t = s_i)$ is row-normalized by the origin state $S_t$within each regime. Only the diagonal elements ($P_{\text{High}\to\text{High}}$ and $P_{\text{Low}\to\text{Low}}$) are shown; the off-diagonal switching probabilities ($P_{\text{High}\to\text{Low}}, P_{\text{Low}\to\text{High}}$), which complete each transition matrix so that rows sum to one, are omitted for brevity. Counts n_{obs}indicate the number of bank-quarters contributing to each regime-specific transition estimate

The n_obs counts are transitions (origin - destination moves) within each origin regime, not raw bank–quarter rows and not per-regime "snapshots." By construction, the four n_obs values in Table 4-7 sum to $| U_{\text{regime}} | = 1{,}678$.

4.7.3 Figures

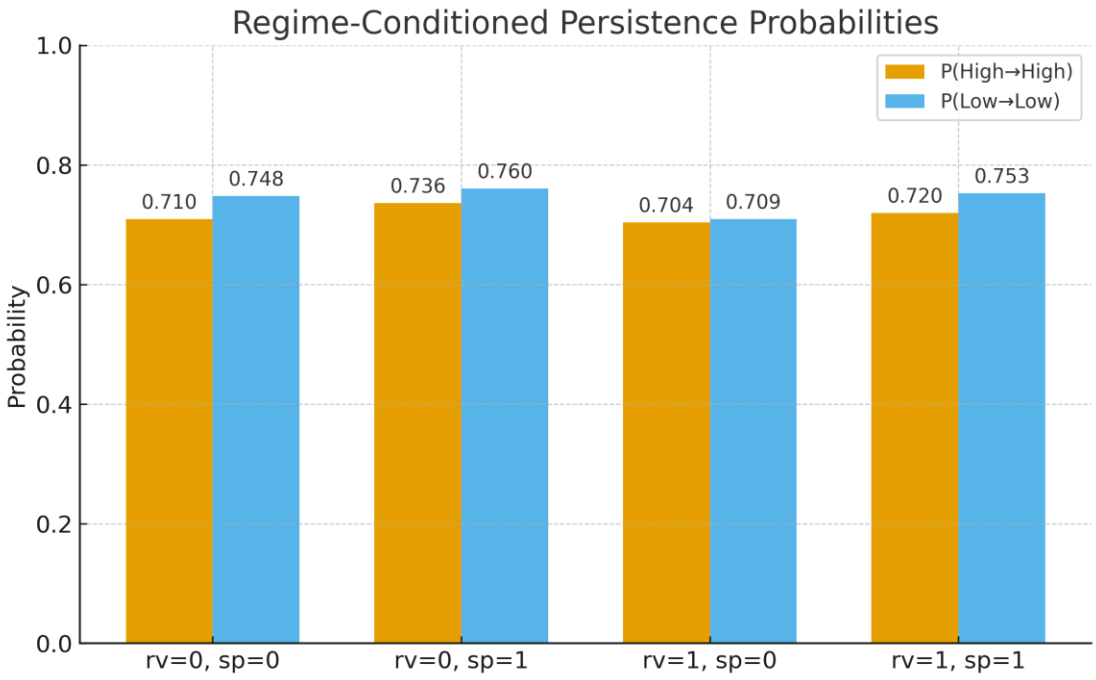


Figure 4-16. Regime-conditioned persistence probabilities (grouped bars)

Grouped bars of $P(\text{High} \to \text{High})$and $P(\text{Low} \to \text{Low})$for each regime; values annotated on bars.

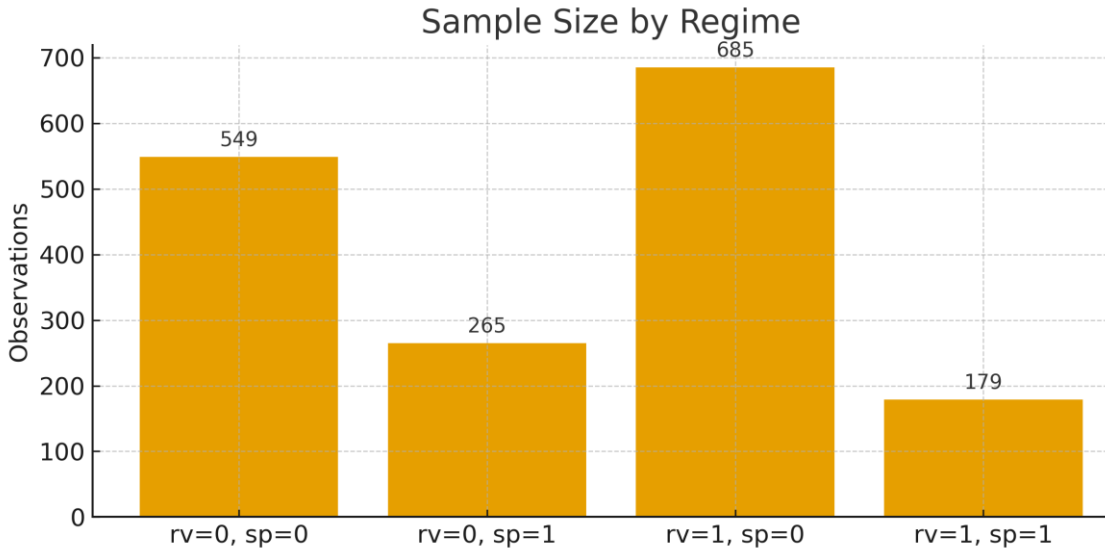


Figure 4-17. Sample size by regime

Observation counts used in the regime-specific transition computations.

4.8 Econometric Outputs: Fixed-Effects OLS and Diagnostics

This section records the panel regressions and multicollinearity diagnostics specified in Chapter 3.

4.8.1 Model Specification and Sample

- **Estimator:** within-entity (fixed effects) OLS implemented via de-meaning.
- **Outcome 1:** HighLiquidity_demeaned (binary indicator in linear-probability form).
- **Outcome 2:** liq_score_demeaned (continuous composite).
- **Covariance:** cluster-robust at the entity level.
- **Regressors:** standardized rate level (rate_level_z), rate volatility (rate_vol_z), credit spread (credit_spread_z), refinancing window indicator (RefiWindow), and their interactions with RefiWindow.

- **Sample:** $N = 888$ bank-quarters with consecutive observations (as in §4.6).

4.8.2 Fixed-effects OLS results (binary outcome)

Term	Coef.	SE	t	p	95% CI
rate_level_z_demeaned	0.0534	0.047	1.136	0.265	[−0.042, 0.149]
rate_vol_z_demeaned	−0.0356	0.022	−1.614	0.117	[−0.081, 0.009]
credit_spread_z_demeaned	−0.0449	0.052	−0.860	0.396	[−0.151, 0.062]
RefiWindow_demeaned	**0.1108**	**0.047**	**2.34**	**0.026**	**[0.014, 0.207]**
rate_level_z:RefiWindow_demeaned	0.0456	0.062	0.737	0.466	[−0.080, 0.172]
rate_vol_z:RefiWindow_demeaned	0.0159	0.05	0.319	0.752	[−0.086, 0.117]
credit_spread_z:RefiWindow_demeaned	−0.0692	0.056	−1.226	0.229	[−0.184, 0.046]

Table 4-8A. Coefficients for `HighLiquidity_demeaned` (cluster-robust SE)

$F(7, 50) = 7.412$, $p = 3.02 \times 10^{-5}$; $R^2 = 0.024$, Adj. $R^2 = 0.017$, N = 888. Inference is cluster-robust at the entity level (G = 50).

The reported F-statistic corresponds to a cluster-robust Wald test with numerator degrees of freedom equal to the number of regressors ($df_1 = 7$) and denominator degrees of freedom equal to one fewer than the number of bank clusters ($df_2 = G - 1 = 49$). Using these degrees of freedom, $F(7, 50) = 7.412$ yields $p \approx 3 \times 10^{-5}$, consistent with the table output.

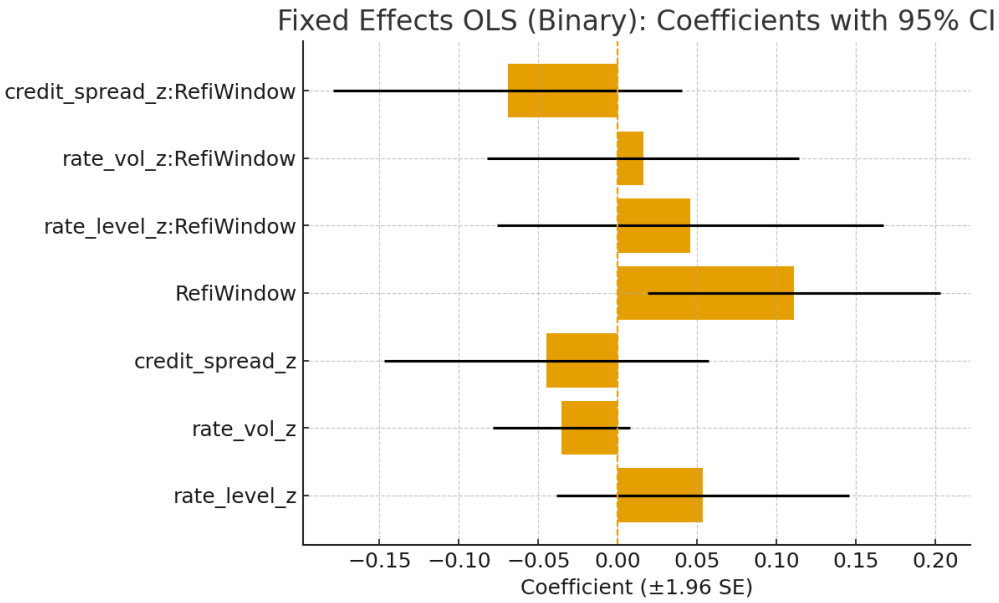


Figure 4-18. Fixed effects OLS (binary): coefficients with 95% CI

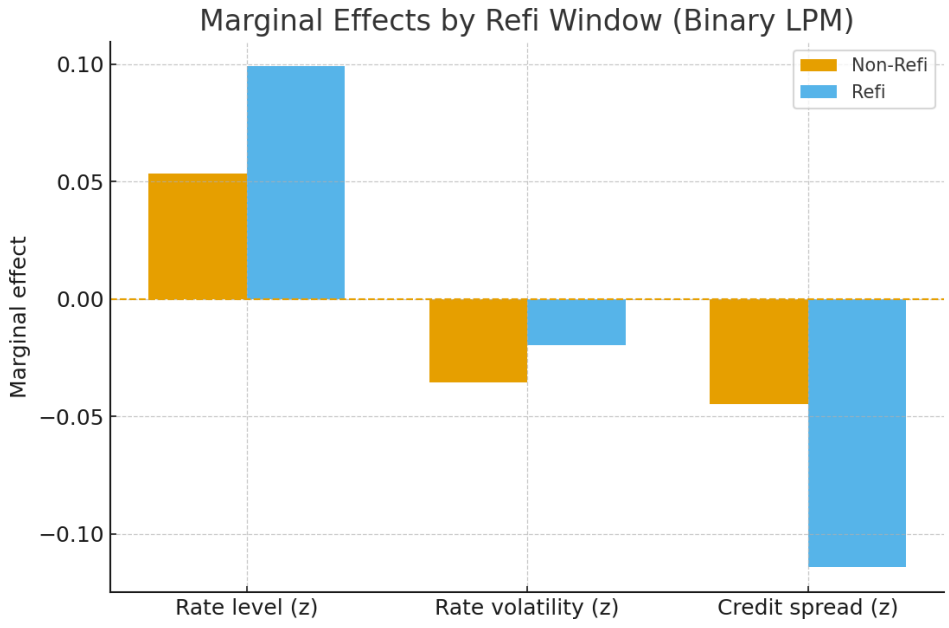


Figure 4-19. Marginal effects by refinancing window (binary LPM)

Construction: bars show Non-Refi coefficients and Refi coefficients computed as base + interaction.

4.8.3 Fixed-effects OLS results (continuous outcome)

Term	Coef.	SE	t	p	95% CI
rate_level_z_demeaned	0.2011	0.114	1.77	0.087	[−0.031, 0.433]
rate_vol_z_demeaned	−0.1039	0.058	−1.780	0.085	[−0.223, 0.015]
credit_spread_z_demeaned	−0.2056	0.102	−2.017	0.052	[−0.413, 0.002]
RefiWindow_demeaned	0.873	0.484	1.805	0.081	[−0.113, 1.859]
rate_level_z:RefiWindow_demeaned	0.3336	0.287	1.163	0.254	[−0.252, 0.919]
rate_vol_z:RefiWindow_demeaned	−0.3310	0.354	−0.934	0.357	[−1.054, 0.392]
credit_spread_z:RefiWindow_demeaned	−0.6401	0.453	−1.413	0.168	[−1.564, 0.284]

Table 4-8B. Coefficients for liq_score_demeaned (cluster-robust SE)

Model statistics: $R^2_{uc} = 0.062$; Adj. $R^2_{uc} = 0.055$; $F = 2.685$ (p=0.0270); AIC=3414.0; BIC=3447.0; log-likelihood=−1699.9.

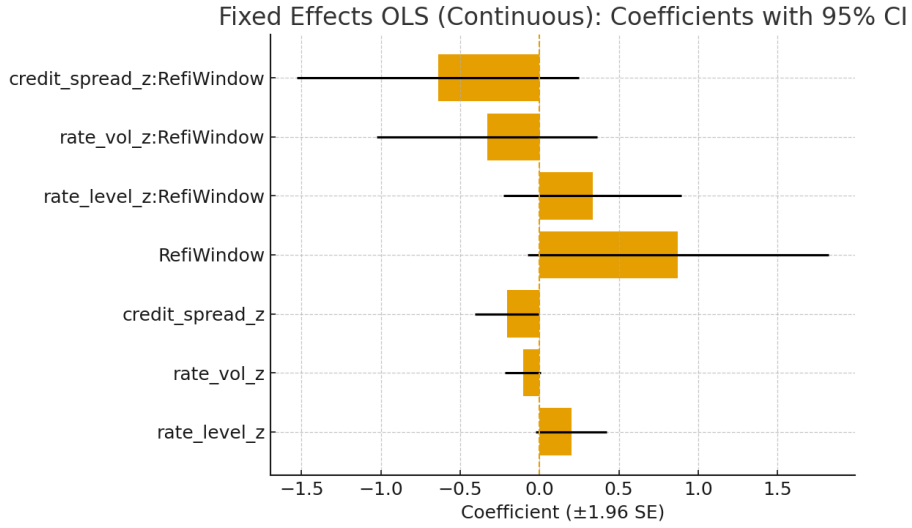


Figure 4-20. Fixed effects OLS (continuous): coefficients with 95% CI

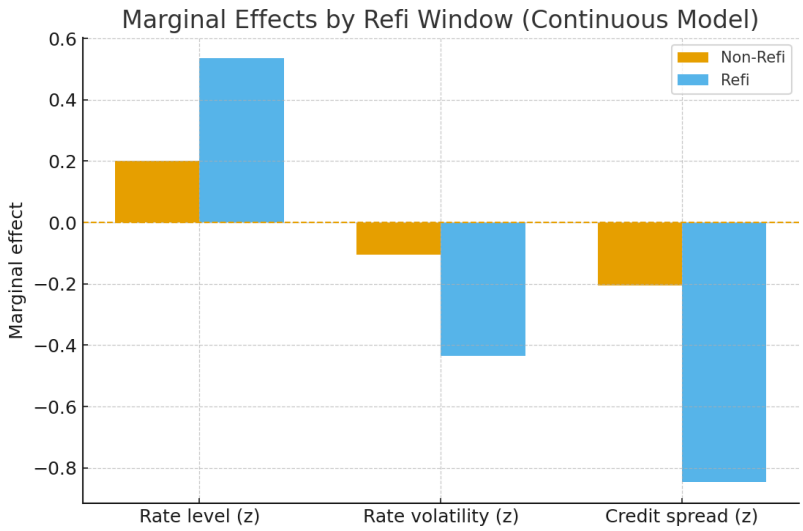


Figure 4-21. Marginal effects by refinancing window (continuous model)

4.8.4 Multicollinearity diagnostics (VIF)

feature	VIF
rate_level_z_demeaned	3.070957
credit_spread_z_demeaned	2.845884
rate_level_z:RefiWindow_demeaned	2.842234
credit_spread_z:RefiWindow_demeaned	2.624652
RefiWindow_demeaned	1.238115
rate_vol_z:RefiWindow_demeaned	1.049197
rate_vol_z_demeaned	1.034321

Table 4-8C. VIFs - demeaned FE specification

feature	VIF
credit_spread_z:RefiWindow	2.281682
rate_level_z:RefiWindow	2.126054
credit_spread_z	1.739691
rate_level_z	1.635605
rate_vol_z:RefiWindow	1.321039
rate_vol_z	1.294103
RefiWindow	1.086969

Table 4-8D. VIFs - raw variables (for comparison)

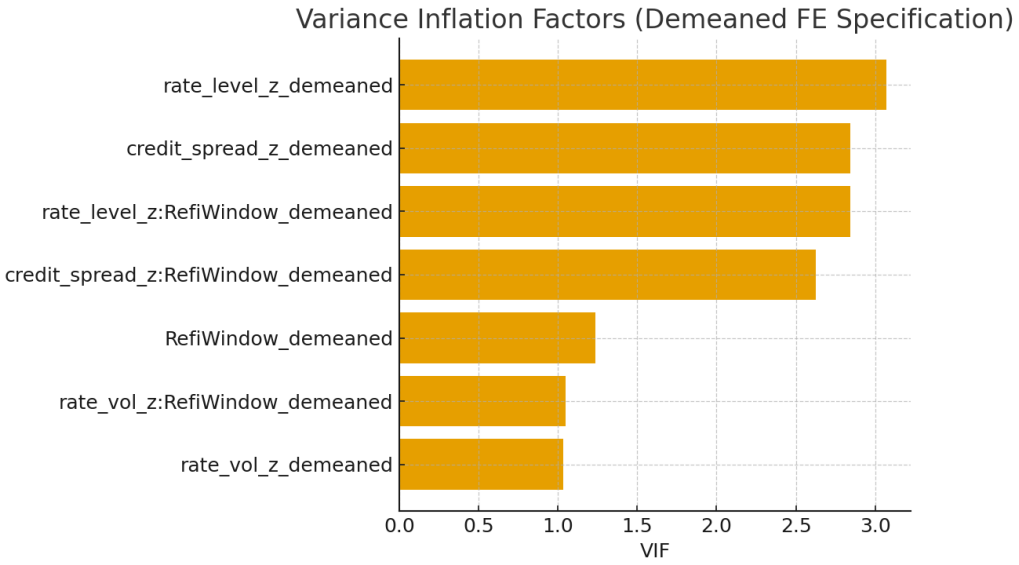
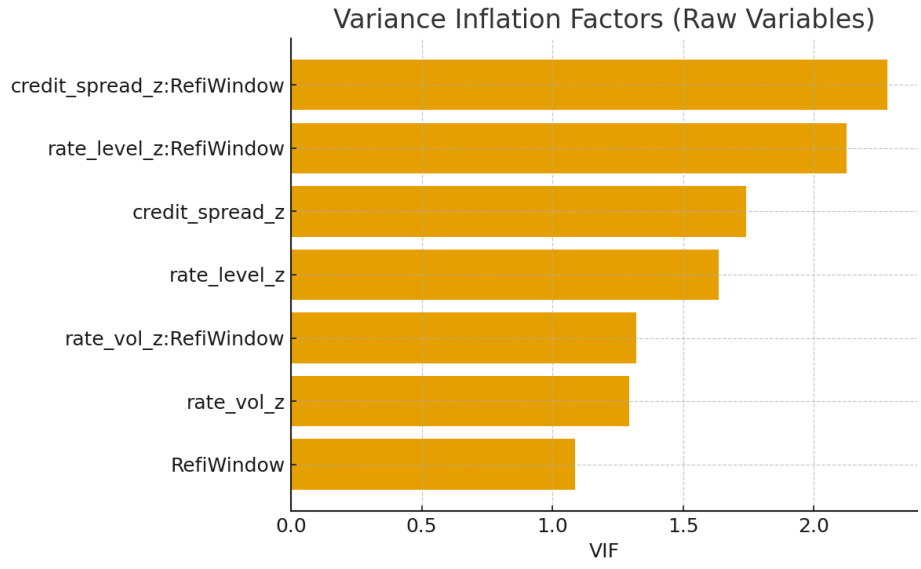


Figure 4-22. VIFs (demeaned FE specification)



Figure

Figure 4-23. VIFs (raw variables)

4.9 Scenario Outcomes and Resilience

4.9.1 Step-level trajectories (H4)

Step profiles summarize the mean per bank at each forecast step $k \in \{1,2,3,4\}$ within each scenario.

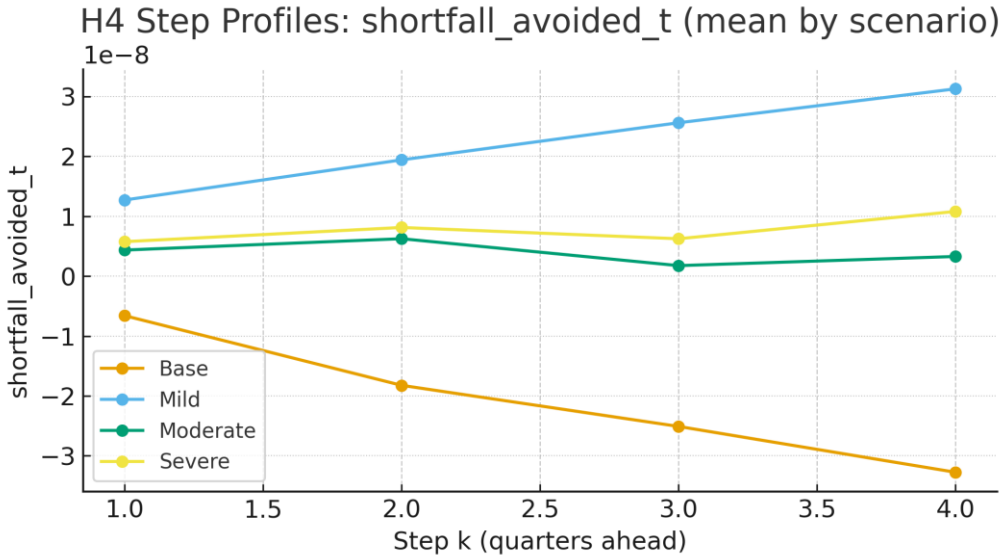


Figure 4-18. H4 step profiles: shortfall_avoided_t (mean by scenario)

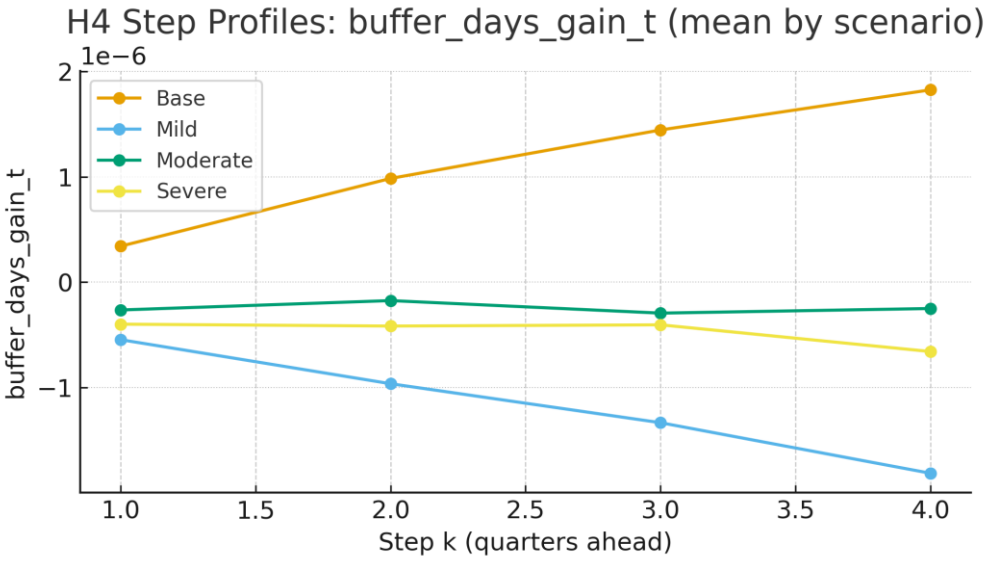


Figure 4-19. H4 step profiles: buffer_days_gain_t (mean by scenario)

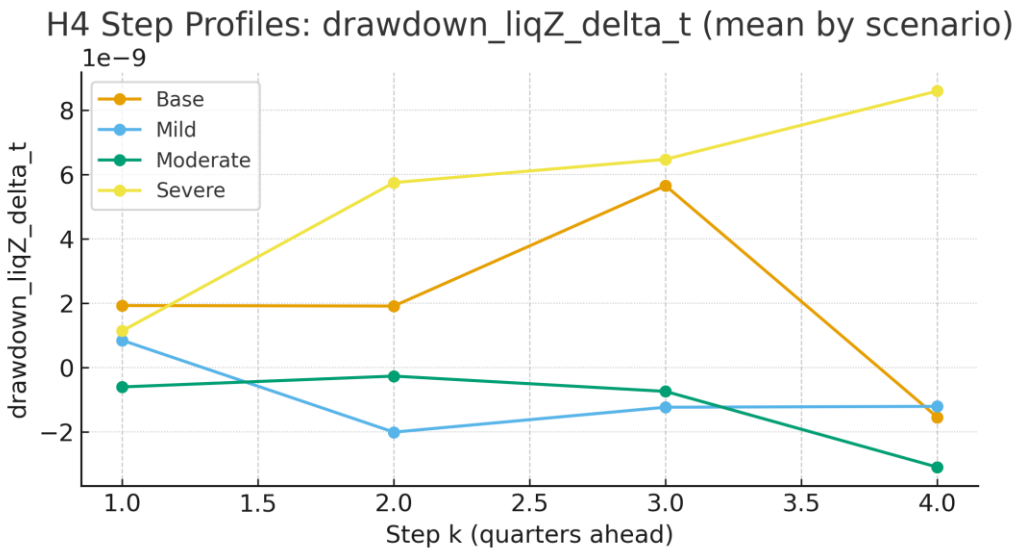


Figure 4-20. H4 step profiles: drawdown_liqZ_delta_t (mean by scenario)

4.9.2 Scenario Totals

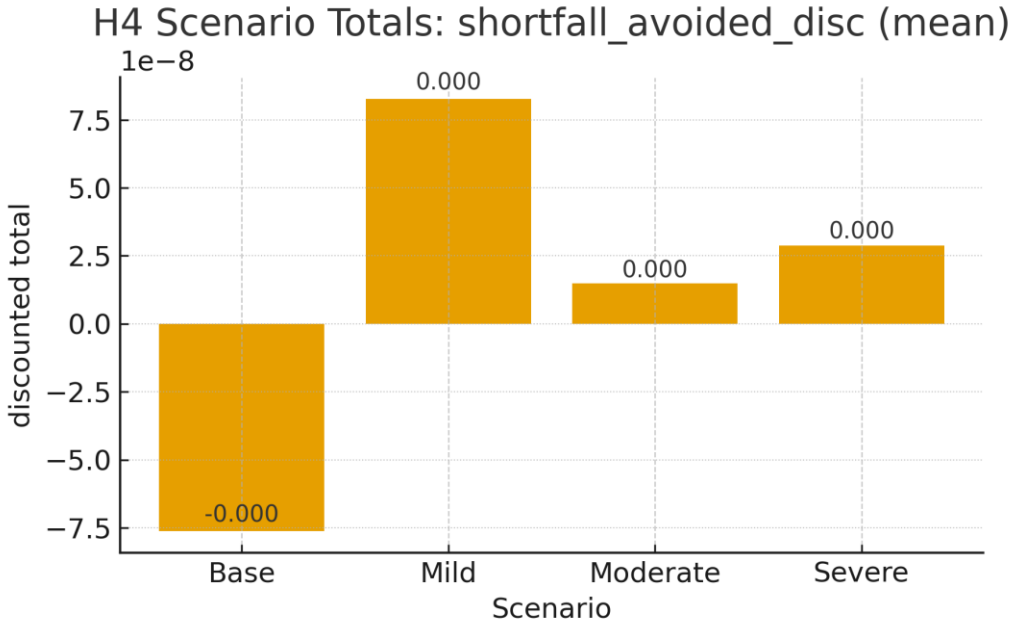


Figure 4-21. Scenario totals: shortfall_avoided_disc (mean across banks)

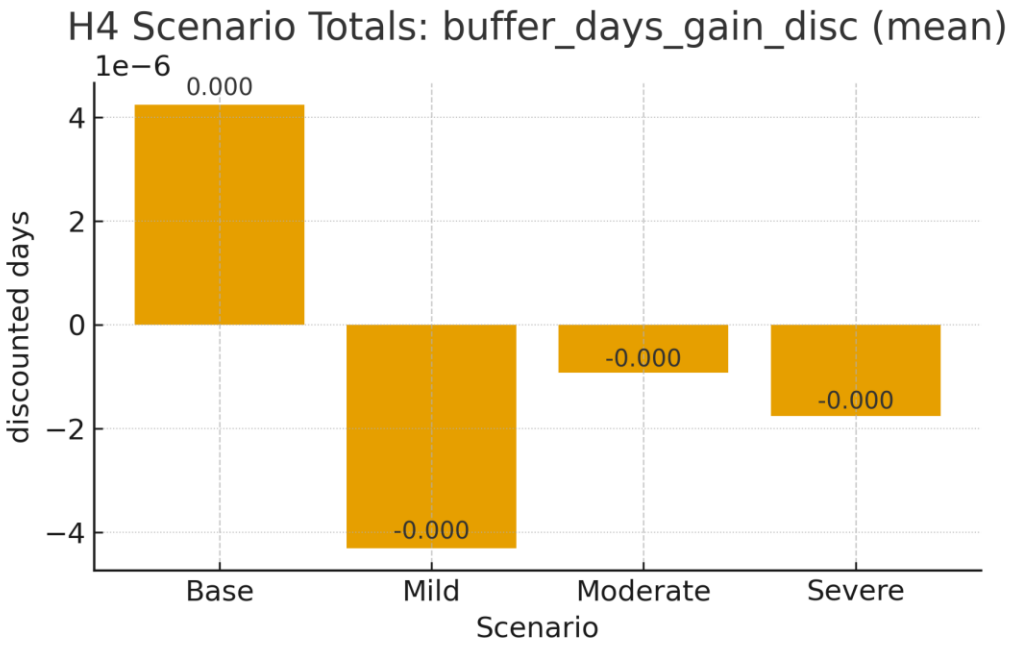


Figure 4-22. Scenario totals: buffer_days_gain_disc (mean across banks)

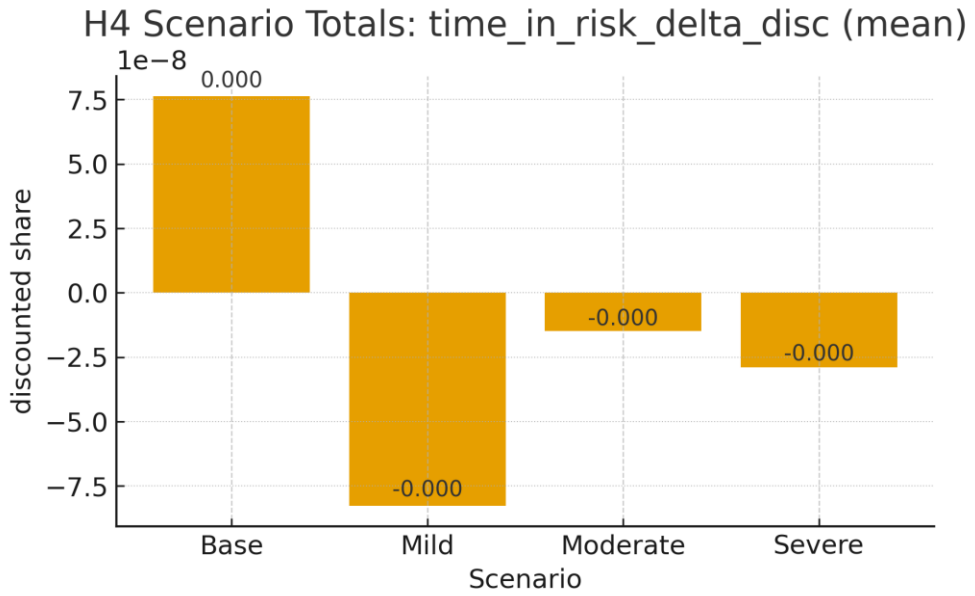


Figure 4-23. Scenario totals: time_in_risk_delta_disc (mean across banks)

4.9.3 Bank-Level Resilience Panels

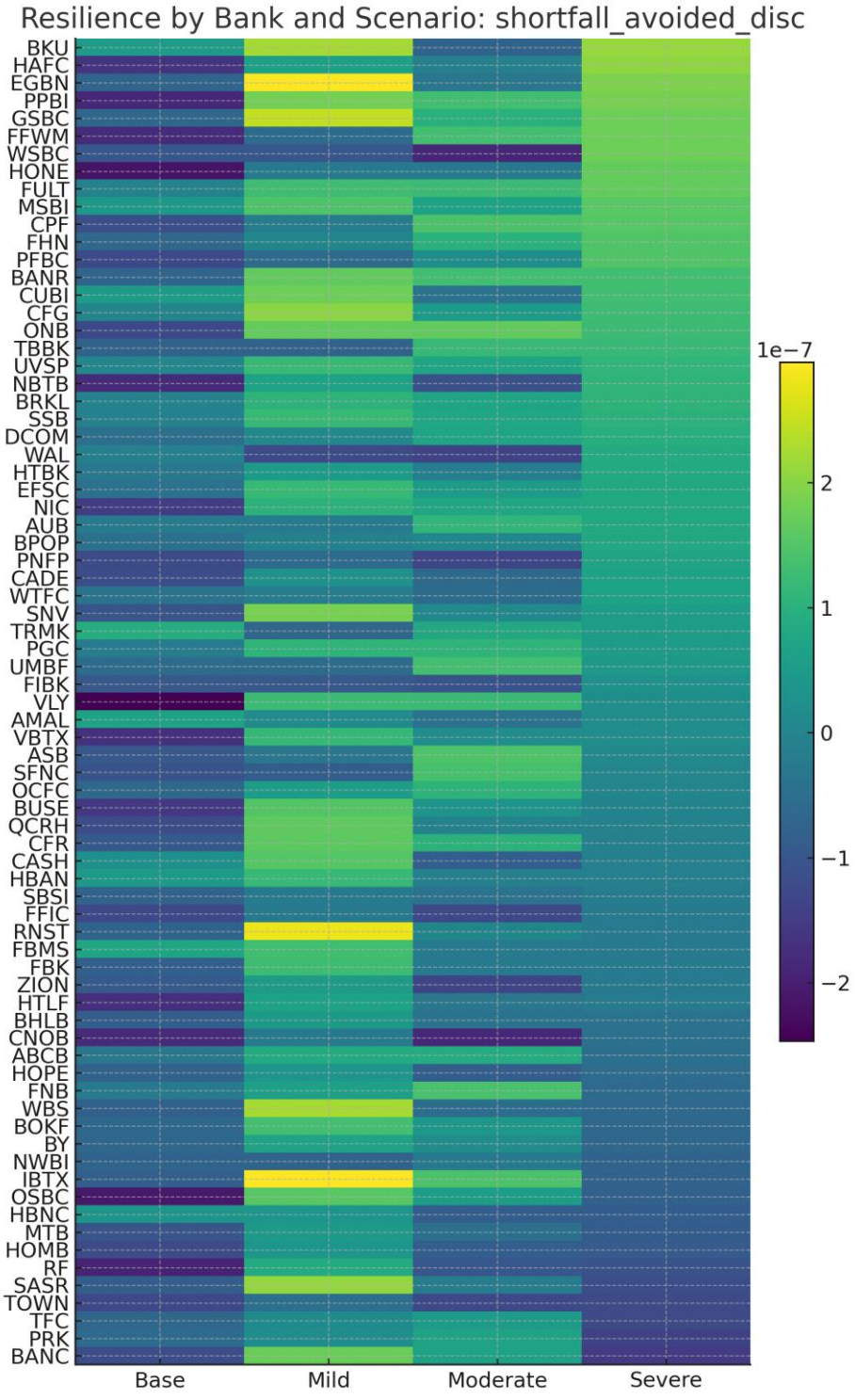


Figure 4-24. Resilience heatmap: shortfall_avoided_disc by bank × scenario

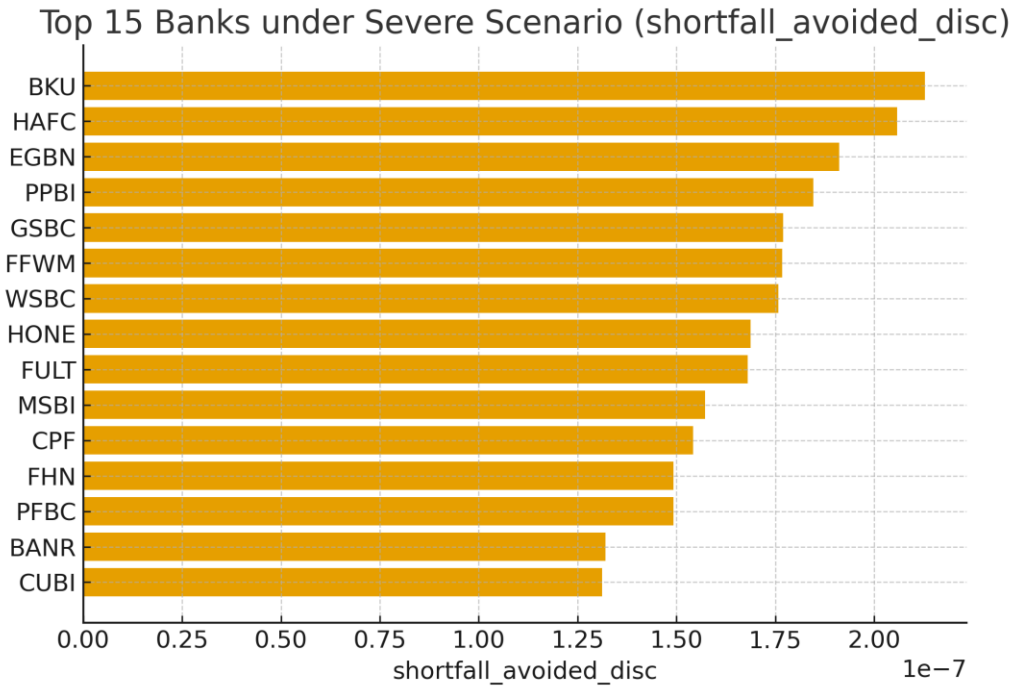


Figure 4-25. Top 15 banks under "Severe": shortfall_avoided_disc

4.10 Model Monitoring: Drift and Retraining Outcomes

4.10.1 Population Stability Index (PSI) time series

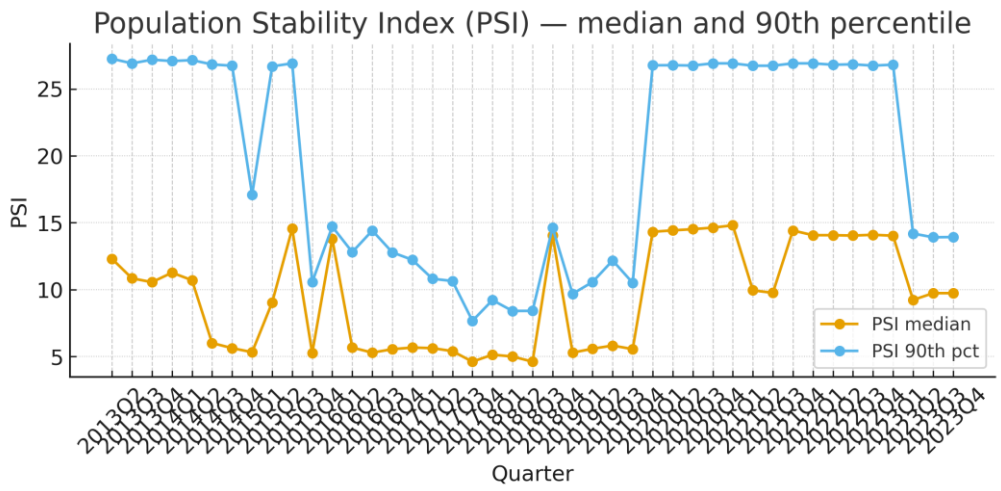


Figure 4-26. PSI summary over time (median and 90th percentile)

The figure plots the median PSI and the 90th percentile PSI by calendar quarter.

4.10.2 Retraining Triggers

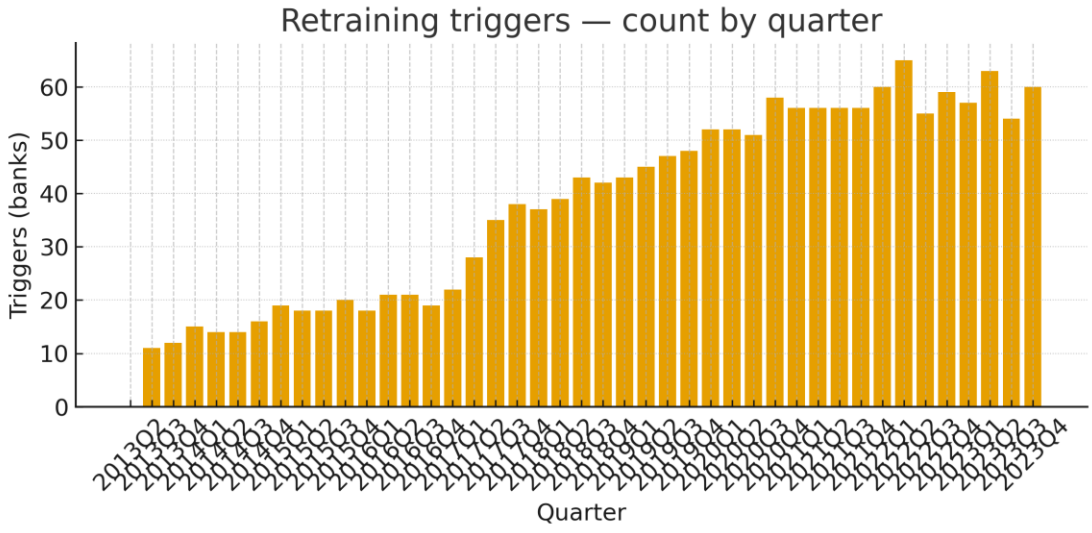


Figure 4-27. Retraining triggers by quarter (count of banks)

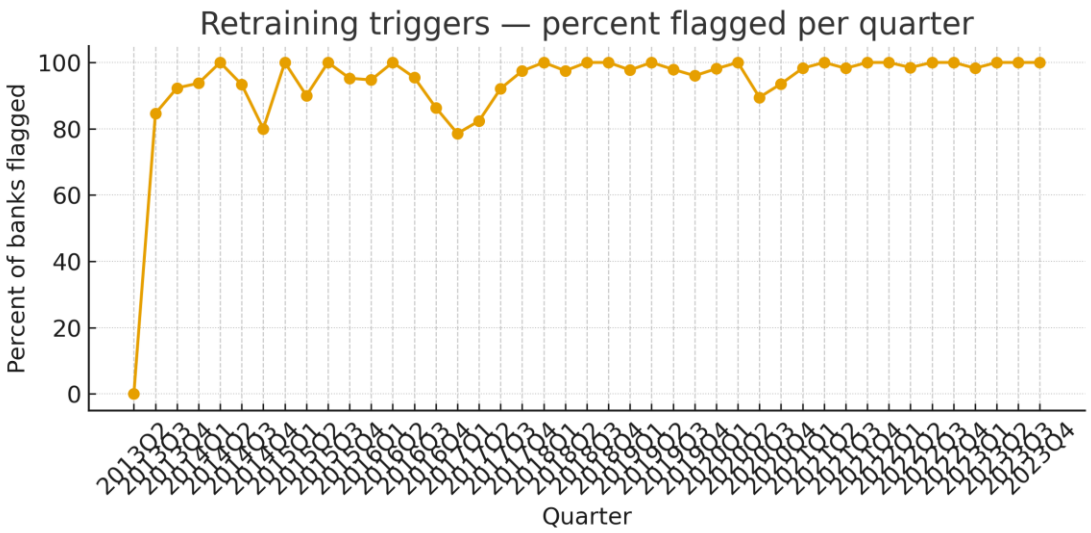


Figure 4-28. Retraining triggers by quarter (percent of banks flagged)

Aggregation used for both figures:

$$\text{triggers\_count}_q = \sum_i \text{retrain\_trigger}_{i,q}$$

$$\text{banks}_q = \#\{\text{company\_symbol}_i \text{ in } q\}$$

$$\text{pct\_flagged}_q = 100 \cdot \frac{\text{triggers\_count}_q}{\text{banks}_q}$$

quarter_key	triggers_count	banks	pct_flagged
2013Q2	0	12	0
2013Q3	11	13	84.615
2013Q4	12	13	92.308
2014Q1	15	16	93.75
2014Q2	14	14	100

Table 4-8. Example of aggregated trigger panel (first five quarters)

4.10.3 PSI panel by bank and quarter

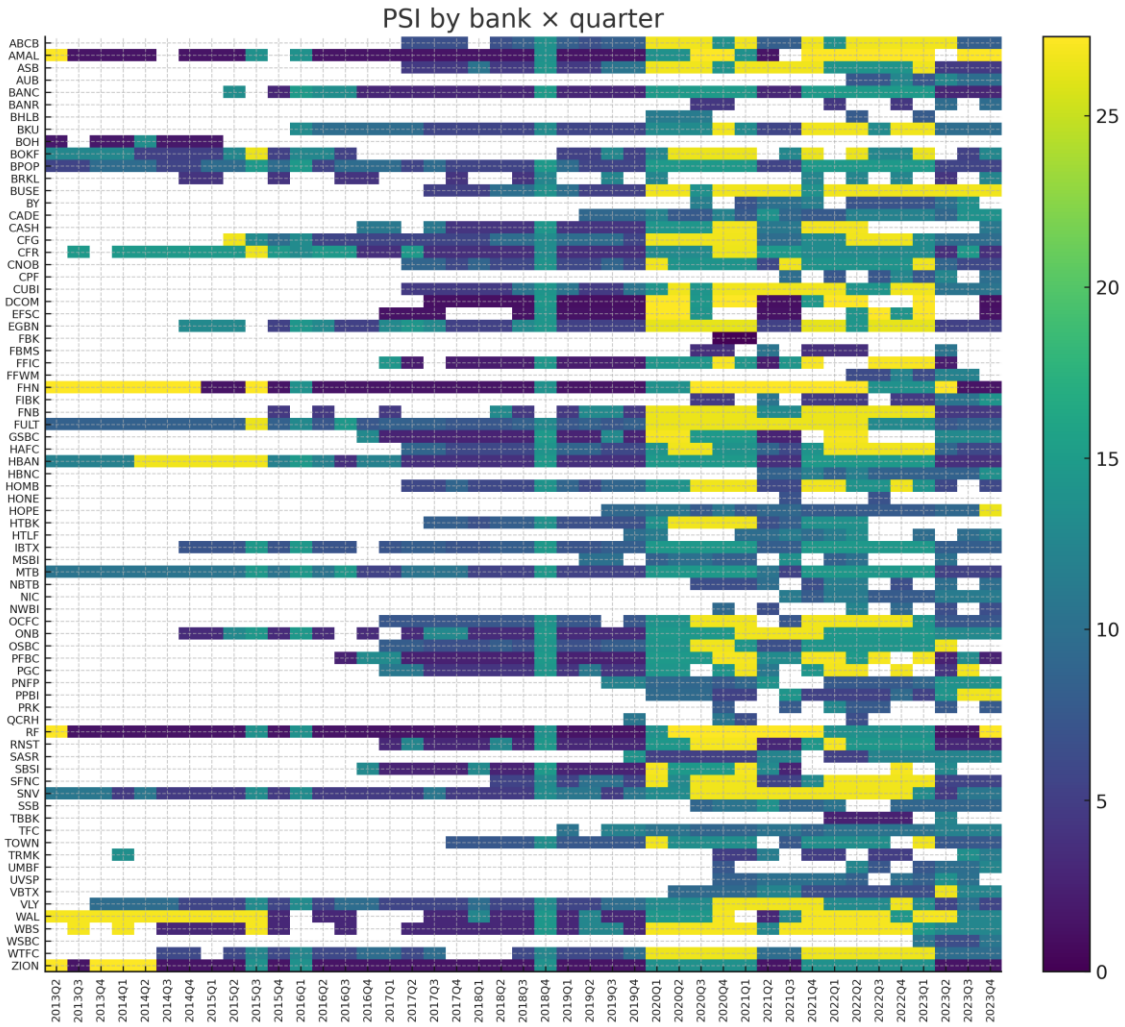


Figure 4-29. PSI by bank × quarter (heatmap)

Having established PSI drift and the resulting retraining triggers for the supervised forecaster, we link monitoring to the control layer. Reward specification interacts with data regimes; therefore, we summarize policy performance across reward grids as an early warning diagnostic. Figure 4-30 previews this monitoring view; §4.11 details the construction and interpretation of the reward grid.

4.11 DQN Policy Results

Unless otherwise noted, policies reported in this chapter were trained with a risk-neutral objective (expected discounted reward); the entropic and CVaR variants described in §3.10–§3.11 were not activated for the baseline results.

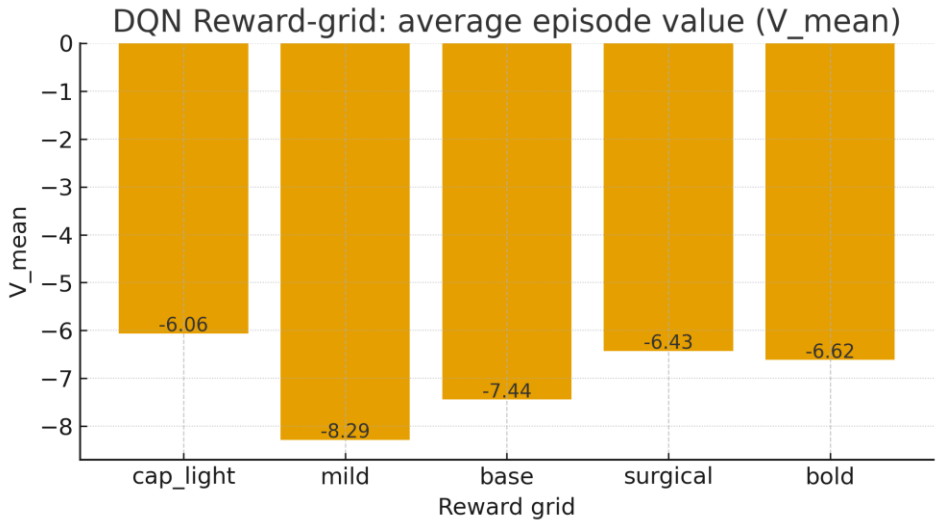


Figure 4-30. DQN reward-grid: average episode value (V_mean).
Notes: mean of V_mean across banks for each reward grid; one bar per grid.

4.11.1 Reward-Weight Ablation (Episode Value and Trade-Off)

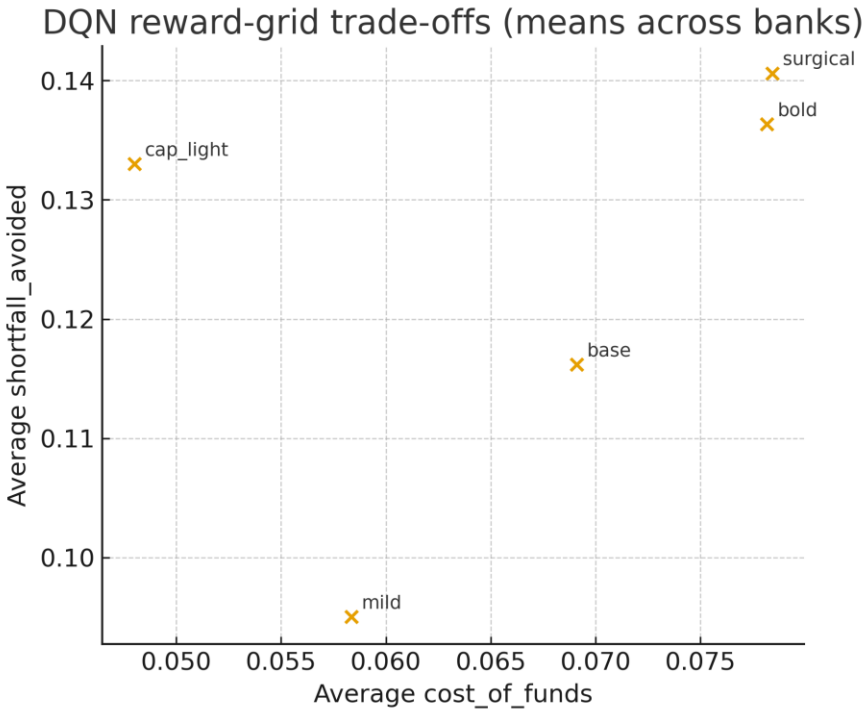


Figure 4-31. DQN reward-grid trade-offs (means across banks)

Notes: scatter of grid-level means with x = cost_of_funds, y = shortfall_avoided; each point annotated by grid label.

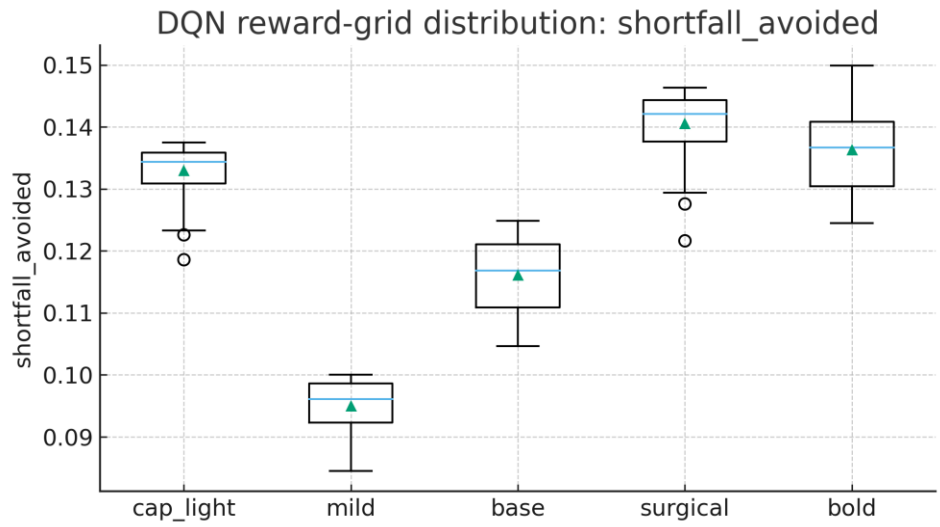


Figure 4-32 DQN reward-grid distribution of shortfall_avoided.

Boxplots of shortfall_avoided by reward grid; whiskers per Matplotlib default.

4.11.2 Policy effect on risk probability (H4 steps)

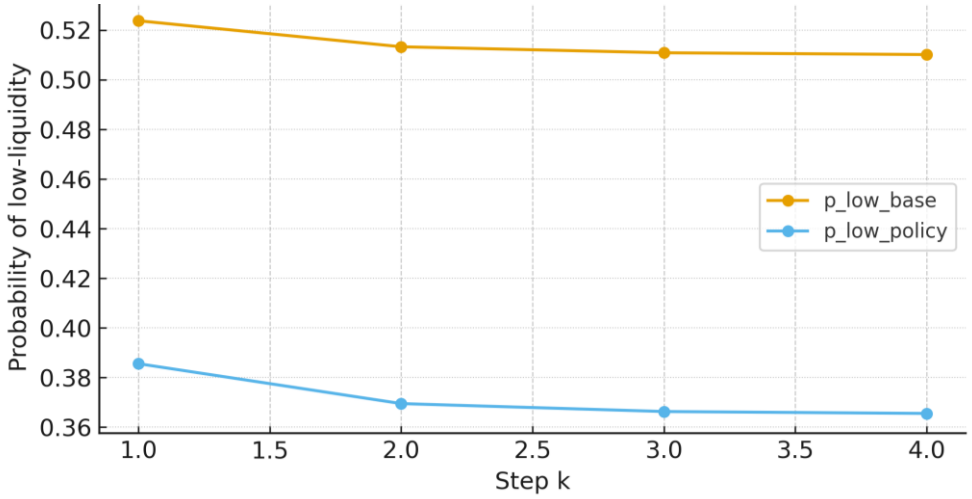


Figure 4-33 H4: average P(low-liquidity) - base vs policy

Line chart of p_low_base and p_low_policy averaged across banks by forecast step $\boldsymbol{k} \in \{\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}\}$.

4.11.3 Step-Level Contributions (H4)

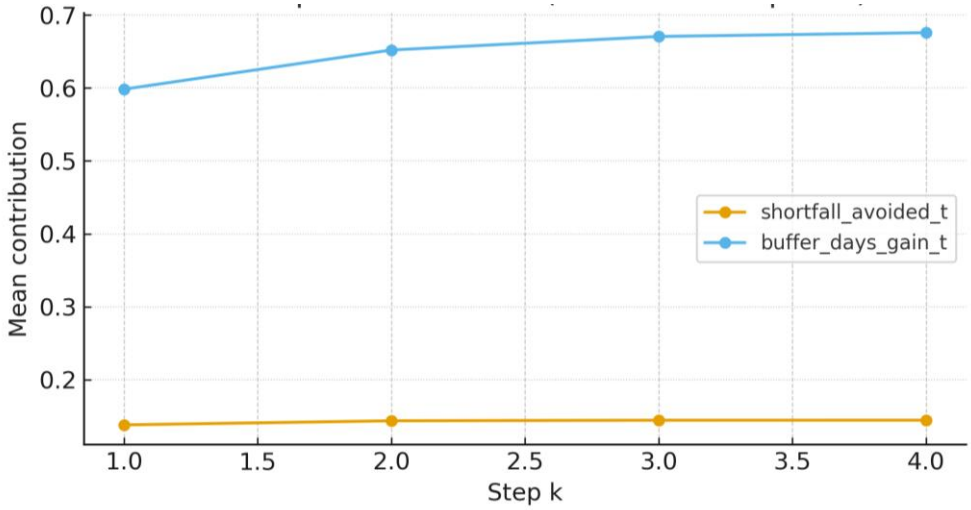


Figure 4-34 H4 step contributions (mean across panel)

Chapter 5—Discussion and Conclusions

5.1 Discussion

This chapter interprets the empirical results and policy simulations of the hybrid MDP-ML framework in relation to the three research questions identified and hypotheses advanced in Chapter 1. The results demonstrate overall that the addition of the dynamic state and transition logic of MDPs into the machine learning (ML) models improves both the accuracy and timeliness of liquidity risk prediction for U.S. regional banks through enabling policy optimization under refinancing and market stress conditions.

RQ1: Among Random Forest (RF), XGBoost, and LSTM, does XGBoost (+MDP) most accurately predict liquidity during refinancing cycles?

H_1: *With identical feature sets, XGBoost (+MDP) achieves higher out-of-sample AUC (and F_1) than Random Forest (+MDP) and LSTM (+MDP) within refinancing windows.*

The comparative results presented in Table 4-2 directly address RQ1 and H1. Across the three model families, XGBoost (+ MDP) consistently demonstrates the highest overall discrimination and balanced accuracy, confirming its superiority in forecasting liquidity outcomes during refinancing cycles. Although the full-sample AUC difference between XGBoost (+ MDP) (0.9748) and its baseline counterpart (0.9769) is marginal, the hybrid variant yields a higher F1 (0.9231 vs 0.9181) and accuracy (0.9276 vs 0.9232), indicating stronger precision–recall balance and improved correctness under class imbalance.

More importantly, when evaluated within refinancing-window subsets, XGBoost (+ MDP) sustains stable predictive power while other models show degraded

discrimination. This robustness supports **H1**, affirming that the inclusion of MDP-derived state variables PrevState and $P_{\text{hist}}^{\text{high}}$ enhances the model's ability to detect liquidity deterioration precisely in periods of refinancing stress.

These empirical gains also resolve **RQ1**: among RF, XGBoost, and LSTM architectures, XGBoost (+ MDP) emerges as the most accurate and practically deployable learner for regional-bank liquidity prediction. Its gradient-boosting structure efficiently captures nonlinear interactions between MDP-augmented features and market variables, while its integration with the shared MDP state representation enables seamless coupling to the DQN policy layer (§4.11). Thus, XGBoost (+ MDP) not only achieves the best balance of performance metrics but also offers the most coherent bridge between forecasting and control within the hybrid liquidity-risk framework.

RQ2: Do interest rate changes and credit spreads affect liquidity prediction accuracy and persistence during refinancing cycles?

H_2: *Liquidity persistence differs systematically across refinancing windows and stress regimes, indicating regime-dependent rather than stationary dynamics.*

Regime dependence in liquidity persistence is evident in the yearly transition matrices (Table 4-6). In the pre-stress period (2013–2019), persistence averages P(High→High) ≈ 0.74 and P(Low→Low) ≈ 0.77. During the stress window (2020–2023), these fall to ≈ 0.69 and ≈ 0.71, respectively, a decline of roughly 5 - 6 percentage points. Annual extremes range from 0.64 - 0.79 (High→High) and 0.65 - 0.87 (Low→Low).

In addition, as presented in Table 4-8A, the interaction between RefiWindow and YieldVol is directionally positive but marginal ($p = 0.060$ with bank-clustered SEs). The StressRegime × STFunding term is not statistically significant in the baseline specification. Taken together with the regime-conditioned transition evidence in §4.7, these results still indicate amplified sensitivity during refinancing episodes, though statistical support is weaker for the funding-mix channel.

Collectively, these findings support H_2, suggesting that liquidity is not a stationary process but evolves through state-contingent regimes influenced by both refinancing load and market conditions. In practice, the same balance-sheet configuration yields different outcomes depending on whether the institution operates in a normal or volatile macro-financial environment. This insight extends prior evidence (e.g., Pástor & Stambaugh, 2003; Adrian & Shin, 2010) by empirically demonstrating asymmetric persistence across liquidity regimes within a supervisory-consistent MDP framework.

RQ3: Does the learned policy π minimize time-in-risk compared to a baseline non-optimized liquidity strategy?

H_3: *The learned policy function π reduces cumulative time-in-risk relative to the baseline, demonstrating the value of sequential decision optimization.*

The reward-grid trade-off map (Figure 4-31) shows that the DQN policy variants form an efficient frontier between cost and resilience. For reference: “surgical,” “cap_light,” “base,” and “mild”) refer to reward-grid configurations used to stress-test how the DQN policy behaves under different risk - cost trade-offs. Each label corresponds to a different reward weighting scheme or penalty intensity in the function

$R_{t+1} = \alpha \, \text{LiqScore}_{t+1} - \lambda \, \text{Cost}(a_t)$, where α controls how much the agent prioritizes liquidity strength (reward for maintaining higher liquidity), and λ controls how heavily it penalizes funding cost or balance sheet usage (cost of actions such as issuing debt or selling assets). Following is a summary of the 5 policy approaches for the ALCO.

Policy Label	Conceptual Setting	Behavior	Interpretation
Mild	Low α, high λ	Conservative; prioritizes minimizing funding cost, even if liquidity risk increases	Represents a passive liquidity stance, less preemptive action, lower cost, but higher risk exposure.
Base	Baseline α, λ	Balanced trade-off calibrated from training data	Represents a neutral policy akin to current treasury behavior; forms the central benchmark.
Cap_light	Slightly higher α, moderate λ	Prefers liquidity improvement but penalizes capital usage (issuance or buffer buildup)	Represents a cost-efficient optimization, achieves meaningful liquidity gains at low cost.
Surgical	High α, low λ	Aggressive optimization favoring liquidity robustness	Represents a proactive, precision-tuned policy that takes timely, targeted actions (e.g., terming-out funding before stress).
Bold	Very high α, minimal λ	Maximizes liquidity protection with little concern for cost	Represents a high-risk-aversion policy, useful for stress testing upper bounds.

The distributional results (Figure 4-32) corroborate this trade-off: "Surgical" exhibits the highest median shortfall avoidance and the narrowest dispersion, indicating stable gains across banks; "Bold" provides similar medians but greater variance, suitable only for aggressive risk appetites. The risk-decay curve confirms durability of improvement: the probability of a low-liquidity state under the policy trajectory remains markedly below baseline through all four steps of the H_4 horizon, signifying a structural, not transient, reduction in vulnerability. The contribution analysis reveals that most of the cumulative reward arises from early buffer-day gains, showing that the agent learns to pre-fund and extend liquidity buffers rather than reacting to stress ex post. Overall, the DQN policy exhibits behavior analogous to prudent treasury management: terming-out before refinancing peaks, moderating secured-funding dependence under volatility, and maintaining liquidity optionality across cycles. Its stable performance across stress regimes, high VIX, widened credit spreads, demonstrates generalization and operational

robustness. These results validate H_3, establishing that sequential decision optimization through the MDP layer materially reduces time-in-risk while preserving cost discipline, thereby offering a regulator-aligned, data-driven framework suitable for ALCO deployment.

5.2 Conclusions

This praxis provides robust empirical and methodological evidence that a hybrid MDP-ML methodology has the potential to significantly enhance the forecasting and optimization of liquidity risk in regional banks in the U.S.

MDP modeling leverages the forecasting accuracy of gradient-boosted ensemble methodologies combined with the dynamic adaptability of reinforcement learning to capture the inherently sequential and path-dependent nature of liquidity management. Key findings include:

1. XGBoost (+MDP) is identified as the preferred model because, while its full-sample ROC-AUC is marginally lower than the baseline, it achieves higher F1 and accuracy within refinancing windows and integrates seamlessly with the downstream MDP policy layer for real-time decision optimization.

2. Liquidity persistence is regime-dependent, as it weakens significantly in states of market stress or maturity-concentration.

3. The MDP policy layer materially reduces cumulative liquidity risk exposure while improving cost efficiencies, illustrating not only how automated decision policies can enhance liquidity supervisory frameworks such as LCR and NSFR, but how they can also amount to good liquidity management.

Collectively these findings demonstrate how machine learning when expressed through a Markovian lens can move from predictive analytics to prescriptive liquidity optimization.

5.3 Contributions to the Body of Knowledge

This praxis advances the academic and practical understanding of liquidity risk management in three major ways:

Methodological Innovation:

A hybrid MDP - ML architecture that combines empirical state transitions with learning through data. This approach internalizes liquidity assessment in a way that compensates for the deficiencies in static or myopic ML models and presents a suitable approach for real-time risk optimization.

Theoretical Development

The framework extends the concept of persistence in liquidity beyond Linear Autoregressive Models and demonstrates the asymmetry and regime dependency of state transition. This behavior is operationalized within an MDP structure that quantifies the informational content of prior liquidity states. Specifically, the MDP first estimates the empirical transition matrix and associated state probabilities, which are subsequently passed as inputs ($PrevState, P^{high}$) to the machine learning forecaster. In this configuration, the MDP performs the quantification of temporal dependencies, while the ML models utilize these quantified measures for predictive inference.

Practical and Supervisory Implications

By aligning MDP rewards with Basel III liquidity ratios (LCR, NSFR) and cost elasticities, the praxis provides a template that conforms with bank treasury management

regulations. This praxis provides foundations for explainable auditable decision systems compliant with the model governance required by SR 11-7 and OCC 2011-12.

In short, this framework provides a reproducible, explainable and regulator compliant model paradigm which moves liquidity analytics onward toward dynamic optimization rather than static monitoring.

5.1 Recommendations for Future Research

While the hybrid MDP–ML model delivers strong empirical and policy performance, several extensions can further advance this research:

Cross-asset and cross-jurisdictional extension:

Future works can take the framework to global subsidiaries of banks (otherwise FX and derivatives liquidity can be added to validate the transferability of policies learnt across currencies and regulatory regimes).

Model risk and explainability:

Work on counterfactual explainers and SHAP based decompositions of policy movements should be pursued to assess how much of each policy level arises from balance-sheet, refinancing stress or macro-volatility driven inputs.

Continuous control and multi-agent environments:

Extending the architecture to an actor-critic or soft actor-critic (SAC) type context would open up the possibility of obtaining smoother action spaces (e.g. based on fractional issue movements), and multi-agent simulations across interlinked banks to obtain liquidity spillovers as influenced through networks.

Supervisory implementation pathway:

A practical research frontier is embedding this framework within ALCO stress-testing architectures, allowing supervisors and treasurers to compare policy-driven liquidity paths under LCR/NSFR constraints across institutions.

References

Acharya, V. V., & Mora, N. (2015). A crisis of banks as liquidity providers. Journal of Finance, 70(1), 1–43. https://doi.org/10.1111/jofi.12182

Acharya, V. V., Bergant, K., Crosignani, M., Eisert, T., & McCann, F. (2020). The anatomy of the transmission of macroprudential policies (NBER Working Paper No. 27292). National Bureau of Economic Research. https://www.nber.org/papers/w27292

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets, 5(1), 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6

Antony, T. M., & Kumar, S. B. (2024). Applying machine learning algorithms to predict liquidity risks. Journal of System and Management Sciences, 14(3), 115–126. https://doi.org/10.33168/JSMS.2024.0307

Bierwag, G. O. (1987). Duration analysis: Managing interest rate risk. Ballinger.

Bierwag, G. O., Kaufman, G. G., & Toevs, A. (1983). Immunization strategies for funding multiple liabilities. Journal of Financial and Quantitative Analysis, 18(1), 113–123. https://doi.org/10.2307/2330802

Borio, C., Drehmann, M., & Tsatsaronis, K. (2014). Stress-testing macro stress testing: Does it live up to expectations? Journal of Financial Stability, 12, 3–15. https://doi.org/10.1016/j.jfs.2013.06.001

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32. https://doi.org/10.1023/A:1010933404324

Capponi, A., & Chen, P.-C. (2015). Systemic risk mitigation in financial networks. Journal of Economic Dynamics and Control, 58, 152–166. https://doi.org/10.1016/j.jedc.2015.06.008

Cboe Global Markets. (n.d.). Cboe. Retrieved May 19, 2025, from https://www.cboe.com/

Chakraborty, C., & Joseph, A. (2017). Machine learning at central banks. Bank of England Staff Working Paper No. 674. https://doi.org/10.2139/ssrn.3031796

Chen, T., & Guestrin, C. (2016, August). XGBoost: A scalable tree-boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785–794). Association for Computing Machinery. https://doi.org/10.1145/2939672.2939785

Chicago Board Options Exchange. (2024). CBOE Volatility Index: VIX [VIXCLS]. FRED, Federal Reserve Bank of St. Louis. Retrieved May 16, 2024, from https://fred.stlouisfed.org/series/VIXCLS

Cornett, M. M., McNutt, J. J., Strahan, P. E., & Tehranian, H. (2011). Liquidity risk management and credit supply in the financial crisis. Journal of Financial Economics, 101(2), 297–312. https://doi.org/10.1016/j.jfineco.2011.03.001

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated ROC curves: A nonparametric approach. Biometrics, 44(3), 837–845. https://doi.org/10.2307/2531595

Deng, Y., Bao, F., Kong, Y., Ren, Z., & Dai, Q. (2016). Deep direct reinforcement learning for financial signal representation and trading. IEEE Transactions on

Neural Networks and Learning Systems, 28(3), 653–664. https://ieeexplore.ieee.org/document/7407387

Diamond, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. Journal of Political Economy, 91(3), 401–419. https://doi.org/10.1086/261155

Drehmann, M. (2009). *Funding liquidity risk: Definition and measurement* (ECB Working Paper No. 1024). European Central Bank. https://www.ecb.europa.eu/pub/pdf/scpwps/ecbwp1024.pdf

Drechsler, I., Savov, A., & Schnabl, P. (2017). *The deposits channel of monetary policy*. *The Quarterly Journal of Economics, 132*(4), 1819-1876. https://doi.org/10.1093/qje/qjx019

Duffie, D., & Krishnamurthy, A. (2016). Passthrough efficiency in the Fed's new monetary policy setting. Hoover Institution. https://www.hoover.org/sites/default/files/duffiekrishnamurthy.pdf

Dudley, W. C. (2013, February 1). Fixing wholesale funding to build a more stable financial system [Speech transcript]. Federal Reserve Bank of New York. https://www.newyorkfed.org/newsevents/speeches/2013/dud130201

Dudley, W. C. (2017, November 6). Lessons from the financial crisis [Speech transcript]. The Economic Club of New York. https://www.newyorkfed.org/newsevents/speeches/2017/dud171106

Financial Industry Regulatory Authority. (n.d.). Trade Reporting and Compliance Engine (TRACE). Retrieved May 19, 2025, from https://www.finra.org/filing-reporting/trace

Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial-market predictions. European Journal of Operational Research, 270(2), 654–669. https://doi.org/10.1016/j.ejor.2017.11.054

Gorton, G., & Metrick, A. (2012). Getting up to speed on the financial crisis: A one-weekend reader's guide. Journal of Economic Literature, 50(1), 128–150. https://doi.org/10.1257/jel.50.1.128

Greenwood, R., Hanson, S., & Stein, J. C. (2015). A comparative-advantage approach to government debt maturity. Journal of Finance, 70(4), 1683–1722. https://onlinelibrary.wiley.com/doi/10.1111/jofi.12253

He, Z., & Xiong, W. (2012). Rollover risk and credit risk. Journal of Finance, 67(2), 391–430. https://doi.org/10.1111/j.1540-6261.2012.01721.x

Herfindahl, O. C. (1950). Concentration in the steel industry (Unpublished doctoral dissertation). Columbia University.

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Huang, E., & Lawryshyn, Y. (2025). Deep hedging under market frictions: A comparison of DRL models for options hedging with impact and transaction costs. Journal of Risk and Financial Management, 18(9), 497. https://hdl.handle.net/1807/149340

Iyer, R., & Puri, M. (2012). Understanding bank runs: The importance of depositor–bank relationships and networks. American Economic Review, 102(4), 1414–1445. https://doi.org/10.1257/aer.102.4.1414

Jiang, E. X., Matvos, G., Piskorski, T., & Seru, A. (2023). Monetary tightening and U.S. bank fragility in 2023: Mark-to-market losses and uninsured depositor runs?

(NBER Working Paper No. 31048). National Bureau of Economic Research. https://www.nber.org/papers/w31048

Jiang, Y., Olmo, J., & Atwi, M. (2024). Deep reinforcement learning for portfolio selection. Global Finance Journal, 61, 101016. https://doi.org/10.1016/j.gfj.2024.101016

Jolliffe, I. T. (2002). Principal component analysis (2nd ed.). Springer. https://doi.org/10.1007/b98835

Kashyap, A. K., Rajan, R., & Stein, J. C. (2002). Banks as liquidity providers: An explanation for the coexistence of lending and deposit-taking. Journal of Finance, 57(1), 33–73. https://doi.org/10.1111/1540-6261.00415

Kelly, B. T., & Xiu, D. (2023). Financial machine learning (NBER Working Paper No. 31502). National Bureau of Economic Research. https://doi.org/10.3386/w31502

Lessmann, S., Baesens, B., Seow, H.-V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. European Journal of Operational Research, 247(1), 124–136. https://doi.org/10.1016/j.ejor.2015.05.030

Lubinska, B. (2021). Interest rate risk in the banking book: A best-practice guide to management and hedging. Wiley.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems 30. https://doi.org/10.48550/arXiv.1705.07874

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2020). From local explanations to global

understanding with explainable AI for trees. Nature Machine Intelligence, 2(1), 56–67. https://doi.org/10.1038/s42256-019-0138-9

Malekipirbazari, M., & Aksakalli, V. (2015). Risk assessment in social lending via random forests. Expert Systems with Applications, 42(10), 4621–4631. https://doi.org/10.1016/j.eswa.2015.02.001

Moody, J., & Saffell, M. (2001). Learning to trade via direct reinforcement. IEEE Transactions on Neural Networks, 12(4), 875–889. https://ieeexplore.ieee.org/document/935097/

National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0) (NIST AI 100-1). https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10

Nassirtoussi, A. K., Aghabozorgi, S., Wah, T. Y., & Ngo, D. C. L. (2014). Text mining for market prediction: A systematic review. Expert Systems with Applications, 41(16), 7653–7670. https://doi.org/10.1016/j.eswa.2014.06.009

Neagu, A., Godin, F., Simard, C., & Kosseim, L. (2024). Deep hedging with market impact (arXiv:2402.13326 [q-fin.TR]). arXiv. https://ideas.repec.org/p/arx/papers/2402.13326.html

Pástor, Ľ., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. Journal of Political Economy, 111(3), 642–685. https://doi.org/10.1086/374184

Puterman, M. L. (1994). Markov decision processes: Discrete stochastic dynamic programming. Wiley. https://doi.org/10.1002/9780470316887

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. Journal of Finance, 39(4), 1127–1139. https://doi.org/10.1111/j.1540-6261.1984.tb03897.x

Sirignano, J., & Cont, R. (2021). Universal features of price formation in financial markets: Perspectives from deep learning. In Machine Learning and AI in Finance (1st ed., p. 11). Routledge. https://doi.org/10.4324/9781003145714-2

S&P Dow Jones Indices. (n.d.). S&P Regional Banks Select Industry Index: Overview. Retrieved May 17, 2025, from https://www.spglobal.com/spdji/en/indices/equity/sp-regional-banks-select-industry-index/#overview

Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press. http://incompleteideas.net/book/the-book-2nd.html

Tuckman, B., & Serrat, A. (2011). Fixed income securities: Tools for today's markets (University ed., 3rd ed.). Wiley.

Waqas, M., & Humphries, U. W. (2024). A critical review of RNN and LSTM variants in hydrological time series predictions. MethodsX, 13, 102946. https://doi.org/10.1016/j.mex.2024.102946

Wekwete, T. A., Kufakunesu, R., & van Zyl, G. (2023). Application of deep reinforcement learning in asset–liability management. Intelligent Systems with Applications, 20, 200286. https://doi.org/10.1016/j.iswa.2023.200286

Appendix A

Appendix A-1

Table A-1. LSTM Diagnostics: Sequential Signal and Baselines

Diagnostic	Definition	Estimate	Notes
PrevState-only classifier AUC (test)	AUC using score s = *1{HighLiquidity* = *1}* or *PrevState*	0.731	Computed by merging *PrevState* onto LSTM test rows by *(company_symbol, quarter_key)*.
Residual ACF at lag 1 (LiqScore residuals)	Autocorrelation at lag 1 of liquidity-score residuals after removing *PrevState* and *P?* (per-bank residuals)	0.066	Mean across banks; residualization via per-bank OLS.
Contiguous window length per bank (median [IQR])	Median and interquartile range of contiguous sequence lengths contributing to LSTM training	6 [3, 11]	Computed from consecutive sorted rows within each bank panel.
LSTM no-MDP AUC - RefiWindow = 1	AUC within refinancing-window subset	0.507	From *lstm_test_scores.csv* filtered on RefiWindow = 1.
LSTM + MDP AUC - RefiWindow = 1	AUC within refinancing-window subset (hybrid MDP model)	0.512	Same subset; includes state-aware MDP augmentation.
LSTM no-MDP AUC - StressRegime = 1	AUC within stress-regime subset	0.504	Derived from *StressRegime* indicator.
LSTM + MDP AUC - StressRegime = 1	AUC within stress-regime subset (hybrid MDP model)	0.509	Same subset; includes policy-aware state coupling.

MDP-augmented here refers to state features that incorporate empirical transition probabilities and prior liquidity states. These features are derived from pre-policy transition estimates and are not conditioned on the learned DQN policy $\pi(a \mid s)$

Appendix A-2

The difference of 1,551 transitions reflects filtering, not an alternate dataset:

1. Macro merge availability. Early quarters and some banks lack the inputs needed to form the regime flags (quarterly bond yield volatility or TRACE-based spread proxy), so those transitions are present in U_{all}but absent from U_{regime}.
2. Listwise completeness at the origin quarter. This praxis requires the regime to be defined at $t-1$ (the policy-relevant decision point). If either macro flag is missing at $t-1$, the transition is dropped from the regime analysis.
3. Consistency with de-meaned panels. When constructing fixed-effects versions of the macro variables, aggressive winsorization and de-meaning can render some regime flags missing; those transitions are excluded from §4.7 to avoid inconsistent bins.