Liquidity Classification of Equities Under Stress Using Machine Learning Models: Evidence from Major World Share Indices

Coskun Tarkocin[1,*] and Murat Donduran[2]

[1]*phD Candidate in Economics, Graduate School of Social Sciences, Yildiz Technical University, Istanbul, Turkey.*

[2]*Director of Graduate School of Social Sciences, Yildiz Technical University, Istanbul, Turkey.*

Abstract: The classification of assets based on their liquidity behaviour under stress is a crucial element of bank liquidity stress testing. It is also important to define how financial institutions should fund these assets within the current business model whilst avoiding excessive liquidity risk. This study aims to revisit the liquidity coverage ratio (LCR) assumptions for common equity shares using new data attributes and supervised machine learning models. This research contributes to the literature by providing fresh insight into which characteristics impact share behaviour under liquidity stress. Empirical results suggest sector, share beta, industry, and market capitalisation of the share are contributing factors which help predict shares' liquidity behaviour under stress. This study also finds that the financial, basic materials and energy sectors are more volatile and less liquid under market stress; shares with lower beta show more liquid characteristics, and higher market cap stocks show more liquid behaviour.

JEL Classification: C10, G01, G21, G33.

Keywords: Liquidity Risk, Equities, Liquidity Coverage Ratio (LCR), Machine Learning Models, Ensemble Model, Random Undersampling Algorithm, financial stress.

1. INTRODUCTION

Over the last 30 years, regulations and technological advancements have significantly transformed the banking industry with instantly available multiple products regulated by complex rules. Development in academic literature and practices of risk management did not prevent the financial crisis of 2007–2008, which was considered one of the worst economic downturns since the Great Depression of the 1930s (Bordo, 2010). Major central banks intervened to stop the collapse of the financial system. Subsequently, the Basel Committee on Banking Supervision (BCBS) introduced a new regulatory framework, widely known as the Basel III rules, to minimise future financial crises. Two new metrics were also introduced for liquidity risk measurement: the liquidity coverage ratio (LCR), and the net stable funding ratio (NSFR) (BCBS, 2010).

The BCBS integrated the new metrics into the Basel Framework in January 2013. The short-term liquidity metric, LCR, requires banks to hold enough unencumbered high-quality liquid assets (HQLA) under idiosyncratic and market-wide stress to meet net outflows over the following 30 days. LCR aims to prevent banks from overreliance on short-term financing and provides a regular liquidity stress test for banks. Regulators expect that a bank should survive 30 days using the stock of the unencumbered HQLA, thereby providing management and supervisors sufficient time to take corrective actions (BCBS, 2013). The long-term funding metric, NSFR, requires banks to have funding sources defined as stable based on their balance sheet structure to increase the resilience of the banking sector (BCBS, 2014b). Bonner and Hilbers (2015) assessed the history of the liquidity regulation until 2013 and found the main reason harmonised liquidity regulation such as this was not introduced earlier was a lack of crisis-related supervisory momentum before the 2007–2008 financial crisis—a crisis mainly driven by liquidity problems.

In this study, LCR assumptions for common equity shares[1] will be revisited. New data attributes and models will be employed and the policy implications of these new approaches will be discussed from both a bank and a regulatory perspective. Liquidity classification is important for financial institutions since it will guide how these assets should be funded within the current business model whilst avoiding excessive liquidity risk.

This paper contributes to the literature in three aspects. Existing liquidity metrics were transformed into a binary classification problem, then supervised machine learning (ML) models were used to predict the classification of the shares under stressed conditions. The results and insights gathered will inform the eligibility criteria of common equity shares and will provide a more granular approach to understanding what impacts the behaviour of equities under stress conditions. Additionally, it will open a new research area for fur-

*Address correspondence to this author at the Candidate in Economics, Graduate School of Social Sciences, Yildiz Technical University, Istanbul, Turkey, Corresponding Author, E-mail: ctarkocin@hotmail.com

[1] Level 2B assets, Equities, Shares and Stocks will be used interchangeably through this study.

ther review of the existing Basel Standards with new data and advanced models. To the best of our knowledge, this is the first time that a set of supervised machine learning models, e.g. the ensemble method with random undersampling algorithm, have been used to further the exploration of liquidity stress assumptions embedded in the Basel Standards and the classification of equities under market-wide stress.

This study is organised as follows: Section 1.2 presents the literature review, Section 1.3 provides liquidity risk definitions, measurements and regulatory classification details; Section 1.4 outlines the data selected for this study, and summarises the handling process along with descriptive statistics; Section 1.5 explains the methodology of this study, provides a brief introduction to machine learning models and explains the selection process for the models used in this study, whilst Section 1.6 discusses the results of this study, including any comparisons between results gathered from the application of different models. Section 1.7 outlines the conclusion of the empirical analysis and discusses its policy implications.

2. LITERATURE REVIEW

This study aims to explore the liquidity characteristics of shares under stress conditions and its focus will be on market liquidity risk and its linkage to the classification of liquidity stress assumptions. The relevant literature is discussed in the following order: (1) literature relevant to the LCR assumptions, (2) literature regarding stock market liquidity and how to measure it and (3) literature concerning machine learning applications in liquidity risk management.

The Basel Committee issued guidance on market-based indicators of liquidity to assist supervisors in evaluating the liquidity profile of assets. Although each jurisdiction determines its own HQLA qualifications, common data and tools help maintain consistency across jurisdictions (BCBS, 2014a). Liquidity standards define HQLA under three categories: Level 1 assets, Level 2A assets and Level 2B assets. Level 1 assets are the highest quality assets with 0% haircut; Level 2A is the next highest quality with a 15% haircut, and Level 2B the lowest quality with a 50% haircut. Subjective and objective criteria for asset classification are provided in the standards at a detailed level (BCBS, 2019).

The assumptions of the LCR were implemented across jurisdictions with very minor changes and challenges. Yet Ball (2020) has criticised the assumptions related to retail deposit outflow, loss of secured funding, and collateral calls under derivatives contracts (mainly the variation margin component) and Level 2B assets. These assumptions were revisited using publicly available data regarding the 2008 crisis as a benchmark. A new liquidity stress test was then developed and applied to six major US banks. Ball argued that, based on the revised LCR assumptions, all six US banks would fail within the 30-day liquidity stress period. In the study, all Level 2B equities were assumed illiquid, yet no data or analysis were provided to support this assumption—only subjective expert judgement was used to apply this stress parameter. Furthermore, Ball (2020) also highlighted that there has been little discussion of specific LCR assumptions by academic researchers.

A detailed report on HQLA characteristics was compiled by the European Banking Authority in 2013 (EBA, 2013b). The report aimed to establish uniform definitions of HQLA characteristics by analysing the wide range in liquidity of the financial assets traded in the EU between 1 January 2008 and 30 June 2012, then classifying such assets from a liquidity and credit quality perspective. The report compares and ranks asset liquidity classes and validates operational and subjective principles that were defined in the LCR. Several liquidity measures were calculated for cross-asset analysis in order to rank them; for instance, when analysing equities specifically, sector and issuance size were also investigated. The evidence for the impact of sector attributes on liquidity is mixed, but it is clear larger issuances have better liquidity values. Nonetheless, the report concludes that there is insufficient evidence of market liquidity to classify equities as “assets of high liquidity and credit quality” (EBA, 2013b, p. 24).

One may argue that the materiality of Level 2B assets is small in the liquidity asset buffers (LABs) of major banks using public disclosures in the US. However, as Ball (2020) argued, material amounts relevant to these assets will be found in the secured funding lines of the LCR. The reason for this is, irrelevant of the source of equities in a bank (outright holding or received as collateral), as soon as an asset is posted as collateral it will be encumbered and will not be shown in the bank’s LAB. Instead, the collateral assets will be returned once the secured financing transaction (SFT) has matured.

As discussed above, academic literature regarding LCR assumptions is limited; in contrast, literature about stock market liquidity and what types of measures can be used is quite vast due to better data availability compared to other asset classes (EBA, 2013b). Jones (2002) provided a comprehensive analysis of the US equity market over 100 years and reported a general decline in the bid–ask spreads on Dow Jones stocks, whilst sharp increases were observed during market stresses. There is noted evidence of liquidity measures such as spreads and turnover predicting returns one year in advance; thus, liquidity is an important determinant of conditional expected returns. Amihud (2002) employed an illiquidity measure (ILLIQ)[2] and conducted news tests which showed asset expected returns increasing in illiquidity. Acharya and Pedersen (2005) built an equilibrium asset pricing model with liquidity risk and used the Amihud measure as an illiquidity proxy. Brunnermeir and Pedersen (2008) developed a model to explain empirically documented features of market liquidity, including sudden dry-ups, commonality across securities, its relation to market volatility, its sensitivity to “flight to quality”, and co-movement with the market.

A long list of liquidity metrics[3] can be found in the literature. Kumar and Misra (2015) classified and organised the literature and provided a critical review of the frameworks cur-

[2] Amihud defines ILLIQ as stock absolute return divided by its daily dollar volume. This study will follow Amihud’s definition as one of the liquidity measures to train the machine learning models.

[3] In this study the term ‘liquidity metric’ is used interchangeably with ‘liquidity proxies’ and ‘illiquidity measures’.

rently available for modelling liquidity. They also presented a summary of the low-frequency liquidity proxies, empirical studies on liquidity proxies, liquidity determinants and liquidity patterns. EBA (2013b) discussed the existing literature for different asset classes, including equities, and then divided the literature into two groups of study: the first which measures liquidity itself, and the second which explores the asset pricing implications of liquidity.

EBA (2013b) listed 25 liquidity metrics, applicable to stocks or securities, and examined eight of these to use in uniform distributions of the assets. Kumar and Misra (2015) listed 18 low-frequency liquidity proxies. Marshall, Nguyen and Visaltanachoti (2013) used three transaction cost benchmarks and nine liquidity proxies to investigate which liquidity proxies measure the actual cost of trading in frontier markets. They found that Gibbs, Amihud and Amivest proxies have the highest correlation with the liquidity benchmarks. Sarr and Lybek (2002) reported nine selected liquidity measures for equity markets in the US, Mexico, South Korea, Malaysia and Indonesia, and noted that liquidity measures may send mixed signals during a crisis. Vayanos and Wang (2012) surveyed the theoretical and empirical literature on market liquidity and reported numerous studies of empirical measures of illiquidity. Fong, Holden and Trzcinka (2017) investigated the most accurate liquidity proxies using both low- and high-frequency data, and found that the Amihud measure is one of the best liquidity proxies among the others. Naes, Skjeltorp and Ødegaard (2011) used four liquidity measures to analyse the relation between stock market liquidity and the business cycle.

The third part of the literature relevant to this study is the application of machine learning models to bank risk management or liquidity risk regulation. A subset of artificial intelligence, supervised machine learning models are employed to conduct data experiments in this study. Machine learning models can be grouped into three main categories: supervised, unsupervised and reinforcement learning. Supervised learning models train data based on a given input and output. By contrast, unsupervised learning models analyse data without a given output, and find potential relationships through clustering data. In reinforcement learning models, the aim is to maximise the defined reward for the specific action given (McKinsey&Company, 2021). Further discussion related to the application of machine learning is provided in the Methodology section.

3. LIQUIDITY RISK DEFINITIONS

There are two types of liquidity risks banks may face in stress. The first is funding liquidity risk. This is where a bank does not have sufficient cash or high-quality collateral to cover liabilities (outflows) as they fall due. Typically, this type of risk is triggered by an idiosyncratic stress event. By contrast, the second type of market liquidity risk, is when a financial asset cannot be sold quickly enough or with a large enough price impact; this type of risk is more driven by market-wide stress. The underlying reason for bank liquidity risk is the traditional banking model, wherein short-term liabilities are converted into longer-term loans by maturity transformation. The main mitigation for liquidity risk therefore becomes establishing a stable funding profile, with a second line of defence provided by having sufficient liquid assets to act as a buffer (Farag, Harland, & Nixon, 2013).

The two types of liquidity risk are closely related to each other. For instance, when funding liquidity risk starts to materialise—which may be for numerous reasons, including large deposit outflows—a bank may need to monetise its liquid asset buffer (LAB) to cover outflows under stress. When a decision is made to monetise non-cash collateral, the market liquidity of the asset becomes critical. For this reason, historically LABs have been comprised of high liquidity and credit quality government bonds such as US Treasuries, UK Gilts and Japanese government debt.

The relationship between market liquidity and funding liquidity is not the focus of this study, therefore it will not be discussed in detail. One of the most cited papers in the literature, Brunnermeier and Pedersen (2008), provides a model that links funding liquidity to asset market liquidity. Under certain conditions, destabilised margins can lead to liquidity spirals. Brunnermeier and Pedersen also show that when speculators face capital constraints, they will reduce risky positions, which later results in a reduction in market liquidity. In this instance, prices will be driven more by funding liquidity considerations than movements in fundamentals.

BCBS (2014a) discussed in detail liquidity characteristics, criteria and metrics as part of guidance provided to the Supervisory Authorities. They defined four main characteristics: asset quality, transparency and standardisation, active and sizeable market, and liquidity (market liquidity). EBA (2013a) followed a two-step approach to rank asset classes and identify explanatory characteristics. The first step was to identify a common set of liquidity metrics and aggregate their results. The second step involved testing whether explanatory characteristics could be used to predict liquidity. Following this EBA (2013a), a detailed report of EBA (2013b) was constructed and no specific change was proposed regarding shares treatment in the regulation.

3.1. Liquid Asset Buffer Assumptions and Its Importance in Bank Liquidity Management

Under liquidity stress, banks could face a significant amount of liabilities leaving such as customers withdrawing deposits or market participants not rolling over short-term financing transactions, depending on the characteristics and severity of the stress event. Banks hold high-quality assets in their LAB such as cash, government securities and other monetisable assets to cover these cash outflows. Simply holding cash assets as a LAB would eliminate the risk associated with monetisability, time taken to monetise and asset price impact. However, holding only cash assets would not be the optimal decision since non-cash HQLA may provide higher returns and natural hedge to banking book positions; it may also need to be held as part of client activity. Banks with a large number of reverse repos or financing transactions receive collaterals which can be used in the LAB, given operational and other requirements are satisfied.

The LCR Delegated Act (LCR DA) definitions and assumptions will be used throughout this study to maintain consistency. In the LCR DA, general requirements (Article 7), operational requirements (Article 8) and eligibility criteria

Table 1. Summary of LCR DA General and Operational Requirements for Liquid Assets.

General Requirements for Liquid Assets (LCR DA Article 7)	Operational Requirements for Liquid Assets (LCR DA Article 8)
The assets should be unencumbered.	LAB is appropriately diversified all the time.
The assets shall not have been issued by the credit institution itself.	LAB should be readily accessible during 30 days.
Not issued by credit institution itself	LAB is under control of the liquidity function.
The assets shall not be issued by a financial company.	Credit institutions regularly monetise the LAB to test monetisability.
The value of the assets can be determined by easily available market prices.	The Assets can be hedged subject to the conditions in Article 8.
The assets shall be listed on recognised exchange or tradable via outright sale or via simple repurchase transaction on generally accepted repurchase markets.	Currency denomination of the LAB is consistent with the currency of net liquidity outflows.

(Chapter 2 Article 10 to Article 17) [4] are defined in detail (LCR Delegated Act, 2015).

Regulatory rules define which assets can be used in a LAB and which haircut percentages must be applied, providing consistency across jurisdictions. Financial institutions may have a different view from regulators, but differing definitions can be applied only in internal metrics; regulatory metrics do not give any flexibility around definitions.

The liquidity classification of assets has three main implications for a bank. First, the classification affects which assets and how much of each type of asset can be relied upon for liquidity stress testing purposes. Second, it affects how assets will be incentivised or disincentivised as part of the fund transfer pricing (FTP) framework. Extremely high credit and liquidity quality assets (e.g. US Treasuries, UK Gilts and German Government Bonds) can be funded short-term, whilst long-term lending to clients or illiquid tradable assets (in some cases short-term assets as well, where franchise risk consideration is high) will require longer-term funding. Third, liquidity classification has external pricing implications. When high-quality collateral is provided as part of secured financing transactions, the haircut applied will be lower compared to that of lower quality HQLA or non-HQLA assets. For these reasons, liquidity classification is an important part of liquidity risk management. This study will contribute to this area of study by providing further insight into what characteristics impact shares' behaviour under liquidity stress.

3.2. Regulatory Treatment of the Shares in Liquid Asset Buffer (LAB)

In this study, the focus will be on the eligibility criteria defined in Article 12(c) of the LCR DA. Fulfilling all requirements of these criteria does not mean an unlimited amount of the assets can be held in the LAB. Caps are applied to each asset quality class to control the composition of the LAB. For instance, the LAB can consist of a maximum 15% of the Level 2B assets (LCR Delegated Act, 2015).

To meet the eligibility requirements of the LCR DA, shares must:

- Be part of a major stock index;
- Be denominated in a member state currency, or can be counted up to net stress outflow in that currency;
- Have a proven record of reliable liquidity source in normal and stressed liquidity conditions. This requirement can be met if the price drop is less than 40%, or the increase in haircuts is less than 40 percentage points during a 30-day period of market stress.

In addition to the above eligibility criteria, general and operational requirements must be met for an asset to be deemed liquid. These requirements are summarised in Table **1**.

4. DATA

Individual stocks from the world's largest stock exchanges are used in this study and data are sourced from Bloomberg and Yahoo Finance. Only data from stock exchanges in developed markets have been analysed to avoid mixing with developing or frontier markets; the rationale for this is developed markets are deeper and more active and show high trading volumes even under stress. Sojka (2019) examined the dynamics of low-frequency liquidity measures for developed and emerging markets and evidenced more liquidity offered on the developed market (Bedowska-Sójka, 2019). This decision has been made to allow more focus on share-specific information.

The New York Stock Exchange (NYSE), Nasdaq, Tokyo Stock Exchange, Shanghai Stock Exchange, Hong Kong Stock Exchange, NYSE Euronext (Europe), London Stock Exchange, and Shenzhen Stock Exchange are by far the largest stock exchanges based on market values (Aras, Karaman, & Kazak, 2020). Data from Shanghai and Shenzhen has not been included in this study, since based on the annual FTSE country classification of equity markets study, as of September 2020, Chinese equity markets are not classified as "Developed" but "Secondary Emerging" (FTSE, 2020).

In the second step of data selection, only major indices from these stock exchanges were selected to investigate the relationship with the largest stocks listed on the markets. This selection will also help retain deep and active market characteristics as eliminating criteria. Therefore, this study's focus will be on how specific shares can be classified given trans-

[4] The LCR Delegated Act may be referred to as LCR DA, LCR rules, or EBA LCR.

parent pricing, available market depth, predicted price drop criteria, and Amihud illiquidity measurements.

4.1. Data Transformation and Cleaning

Collecting all shares in the selected major indices initially left 1100 unique share ISINs. Shares that were not available in the global financial crisis of 2007–2008 were removed to leave a list of shares continuously available between January 2007 and June 2019. After removing duplicate shares between different indices, 882 unique shares remained for analysis. Mainly Euronex100, Eurostocks50, CAC40, DAX30 or Nasdaq100 versus SP500 include several shares from both indices. Variables which are continuously available for all remaining shares are kept for the final modelling training and prediction stage.

In order to test liquidity under stress condition, first we need to define the most stressful period in the last 12 years. Table **2** shows the Global Financial Crisis Period 1 as the most stressful event for the global financial markets in this time period, with an average largest monthly price drop of 50%. Machine learning models will be applied for this period to investigate what may define the shares liquidity under this stressed condition.

Table 2. Monthly Price Drop Across Stress and Historical Periods.

Period	Largest Monthly Drop
Global Financial Crisis Period 1 (Sep–Nov 08)	−50%
Global Financial Crisis Period 2 (Jan–Mar 09)	−31%
European Debt Crisis (Mar–Nov 11)	−26%
Last 5 Years (Jun 14–Jun 19)	−25%

For each share information in Table **3** sourced, or calculated to train the models.

5. METHODOLOGY

Mainly due to increased availability of data, computing power and improved software, the popularity of the machine learning models has increased in the financial sector (BOE, 2019).

Odom and Sharda (1990) conducted one of the earliest studies applying machine learning models in bank risk management and showed the applicability of the neural network model for bankruptcy prediction. Chatzis et al. (2018) used deep and statistical machine learning methods to forecast the stock market crisis and found that data classification accuracy significantly improved with the application of these models. Balaji et al. (2018) applied deep learning models for stock price forecasting and generated an accurate forecast of the direction up to 71.95%.

Leo, Sharma and Maddulety (2019) conducted a literature review of machine learning models in banking risk management and reported many areas in which banking could benefit significantly from their application, including liquidity risk. Several studies listed for credit, market and operational risk applications, however, show that the application of machine learning to liquidity risk management is thus far very limited. One example of such research is Tavana et al.'s 2018 study, in which the researchers employed Artificial Neural Network and Bayesian Networks to measure liquidity risk, demonstrating these models' applicability and efficiency for bank liquidity risk management. Another example is Khan et al.'s 2020 study, whereby deep learning models were used to predict Vietnamese stock market liquidity from a sample of 220 companies' daily stock trading data.

(Nosratabadi et al., 2020) conducted a comprehensive review of advanced machine learning and deep learning methods applications in economics. According to this recently published detailed review, machine learning models are used for stock price prediction, algorithmic trading, portfolio management, sentiment analysis, customer behaviour analysis,

Table 3. Monthly Price Drop Across Stress and Historical Periods.

	Feature Name	**Data Type**	**Explanation**
Features (Input Data)	Sector	Categorical	9 sector, Financial, Basic Materials, Energy, Industrial, Consumer-Cyclical, Technology, Communications, Consumer-Non-cyclical, Utilities
	Industry	Categorical	67 Unique sub industry
	Share Beta as of 29Aug 2008	Numerical	Share Beta pre-Lehman Collapse calculated
	Log (Median Market Cap)	Numerical	Natural logarithm of the median of market capitalisation of the share during stress months
	Median Market Cap Percentage	Numerical	Median of market capitalisation share divided all shares total median market capitalisation (882 shares total)
	90-day average trading volume	Numerical	90-day average trading volume calculated for each share (minimum observed during stress period used)
	Response Name	**Data Type**	**Explanation**
Responses	Cumulative Maximum Price Drop (CMPD)	Categorical	Numerical value transformed into class label (Liquid/Illiquid)
	Amihud Measure (Amh)	Categorical	Same as above

dynamic credit risk evaluation and bankruptcy prediction, amongst other uses. Stock price prediction is the most studied area, followed by marketing including customer behaviour analysis, and then corporate bankruptcy (Nosratabadi, et al., 2020).

To the best of our knowledge, this study will be the first in the literature to use supervised machine learning models to understand what characteristics impact the behaviour of equities under stress conditions. Incorporating the findings of this study will have cost implications, both from a regulator and an individual bank perspective. For this reason, any implementation must have controls and monitoring in place, whilst expert judgement should be applied where required.

5.1. Liquidity Measures

In this study, we will employ two liquidity measures. The first measure is the criteria defined by the Basel Committee and other regulators for eliminating shares from inclusion in a LAB. If a share has more than a 40% price drop in a normal or stressed condition, it is classified as illiquid. In this study, cumulative maximum price drop (CMPD) is calculated for the defined stress period. If a share has more than a 40% drop, it is labelled as illiquid in the empirical analysis to train the model.

- Cumulative Maximum Price Drop (CMPD): Monthly log return is calculated and then a minimum of it is used. The minimum for a negative return will result in the maximum price drop, since in a stress period, equity prices will drop significantly. N is defined as the number of business days used to calculate CMPD.

$$CMPD = Min_{Stres\ period}\left(\frac{\ln(r_t)}{\ln(r_{t-N-1})}\right)$$

The second liquidity measure was first proposed by Amihud (2002) and has since been widely examined in the literature. This measure has been selected for this study due to its simplicity and prevalence in the literature. Additionally, it does not require high-frequency data. For Amihud's measure, liquidity is defined as a daily absolute return on the trading volume for each day. Like the CMPD, it will be used to train models, however one limitation is the need to split what would be the threshold for illiquidity. It will be assumed the illiquid portion will be similar to what the CMPD measure proposes.

- The Amihud illiquidity measure (Amihud): The maximum Amihud measure calculated during the selected stress period is used to train machine learning models, as the higher the Amihud measure, the higher the illiquidity behaviour. In the formula below, N is the same number of business days used in the CMPD measure (20 business days). Stock i on day d, D_i is the number of days used. For each day, the Amihud measure is calculated then an average is taken for the given period. The measure is also used with a slight variation to capture behaviour under stress. Amihud is a widely accepted illiquidity measure and simply presents the price impact of dollars traded, as demonstrated below (Amihud, 2002):

$$Amihud_i = Max_{Stress\ period}\left(\frac{1}{D_i}\sum_{t=1}^{D_i}\frac{|r_{id}|}{volume_{id}}\right)$$

MATLAB 2021a version is used for the implementation of the supervised machine learning models which has eleven ensemble learning algorithms. For full details see MATLAB documentation under Ensemble Algorithms (MATLAB, Ensemble Algorithms, 2019).

The main focus of this study will be on ensemble classifiers since the underlying data is imbalanced. By combining predictions from several base estimators, ensemble learning aims to achieve more robust single estimator (scikit-learn, 2020).

As part of the ensemble models, several boosting methods can be used; in this study we will show the superiority of the RUSBoost algorithm which was first introduced by Seiffert et al. (2008) to reduce class imbalance problems in the data set. RUSBoost uses random data sampling with boosting, which, as a result, improves the classification performance of the training data. Financial stress classification problems have imbalanced data, wherein one class has fewer members than others. The RUSBoost algorithm is used for the modelling in this study to show its effectiveness for the imbalanced data. For a comprehensive overview of the RUSBoost algorithm, please refer to Seiffert et al. (2010).

The RUSBoost applies adaptive boosting for multiclass classification when calibrating weights and constructing ensembles. MATLAB uses weighted pseudo-loss for N observation and K classes. Pseudo-loss (ε_t) is a measure of classification accuracy (MATLAB, Ensemble Algorithms, 2019).

$$\varepsilon_t = \frac{1}{2}\sum_{n=1}^{N}\sum_{k\neq y_n}^{K} d^t_{n,k(1-h_t(x_n,y_n)+(h_t(x_n,k)),}$$

- Each step represented by *t; k* represents class; *N* represents number of observations;
- x_n is a vector of predictor values for observation *n*;
- y_n represents the true class value taking one of the *K* values*;*
- h_t represents the prediction of the learner for each step *t*;
- $h_t(x_n, k)$ is the confidence of the learner prediction at step *t,* class *k* ranges from zero to one;
- $d^t_{n,k}$ represents the observation weights of class *k* in step *t*.

5.2. Performance Evaluation Metrics and Definitions

Confusion Matrix

A confusion matrix was constructed to evaluate the performance of the models. Several performance measures were then calculated from the data presented in the confusion matrix.

The table below (Table **4**) summarises the information presented on the Confusion Matrix.

Table 4. Confusion Matrix.

		Share Predicted Class		
Share Actual Class	**Class**	**Liquid**	**Illiquid**	**Total**
	Liquid	True Liquid (TL)	False Illiquid (FI)- Type 1 Error	**Liq**
	Illiquid	False Liquid (FL)- Type 2 Error	True Illiquid (TI)	**Illiq**
	Total	**Liq***	**Illiq***	**N**

Table 5. Performance Evaluation Metrics.

Measure Name	**Formula**	**Description**
Accuracy (Acc)	$\frac{TL+TI}{N}$,	This metric measures how many observations (both liquid and illiquid) were correctly classified by the model.
Error (Err)	$\frac{FL+FI}{N}$,	This metric provides the misclassification percentage.
Sensitivity (Sens)	$\frac{TL}{TL+FI}$,	True Liquid Class Rate. This measures how many shares out of all liquid observations have a model classified as Liquid.
Specificity (Spec)	$\frac{TI}{TI+FL}$,	True Illiquid Class Rate. This measures how many shares out of all illiquid observations have a model classified as Illiquid.
Balanced Accuracy (BA)	$\frac{SE+SP}{2}$	Average of Specificity and Sensitivity measures.
Weighted Balanced Accuracy Liquid (WBA_L)	$\frac{0.75xSE+0.25xSP}{2}$	Weighted Balance Accuracy, where more weight is assigned to the sensitivity in liquidity classification, where predicted Liquid shares are assigned more weight.
Weighted Balanced Accuracy Illiquid (WBA_ILL)	$\frac{0.25xSE+0.75xSP}{2}$	Similar to above, more value is assigned to the predicted Illiquid shares.

N= Total Number of Data Points= TL+FL+FI+TI, Total number of data points, or number of unique shares used in the modelling process.

Formulas for the measures above are outlined above:

6. EMPIRICAL RESULTS

6.1. Implementation of Models and Assumptions

In the MATLAB program (2021a version), a set of classification models was used to train machine learning models and then predict share liquidity classifications. Training the models first required calculating liquidity measures (responses) and defining split criteria. For the price drop criteria, the 40% eligibility criteria as defined by LCR DA regulations was employed as opposed to the Amihud defining threshold, a limitation of which is that classification assignment is subjective. To split shares based on the Amihud measure, approximately the same percentage of liquid/illiquid from the CMPD classification was used.

For each measure, data experiments were performed which include several distinct models. New functionality in the 2021a version of the MATLAB program automatically searches for the best-performing algorithm and hyperparameters if it is optimisable. The results for eight performance evaluation metrics are presented in the comparison tables (Table **9**, Table **10**). Predicting one specific class (liquid or illiquid) may be more important for an institution or researcher, however in this study predicting both classes is assumed to be equally important. For this reason, a model with a high balanced accuracy, where the gap between sensitivity and specificity is also relatively small, would be preferred.

6.2. Results using Cumulative Maximum Price Drop (CMPD) for Model Training

In order to get some initial perspective, a scatter plot of the results can be a useful tool. In Fig. (**1**), market capitalisation (Market Cap) of the share vs share beta is represented. Blue dots represent shares identified as liquid under stress, with a CMPD of less than 40% during a stress period. Visual inspection of the original observations shows that as market capitalisation increases, the blue dots intensify, whereas when share beta is comparatively lower, the top left corner shows more liquid behaviour.

Table **6** summarises the number of shares falling in each class using a CMPD condition of 40%. Overall, 61% of the shares were reported as illiquid. The financial, basic materials and energy sectors showed the highest percentage of illiquid shares, whilst the utilities and consumer, non-cyclical sectors had the most shares classified as a liquid.

Table **7** shows the average log (market cap) across sector and liquidity classes. For all sectors except financial, the liquid class has higher average market capitalisation. This intuitively supports what Fig. (**1**) shows, and supports the fact that market cap can be a useful measure for predicting the liquidity class of shares.

Fig. (1). Original Data Set Market Cap vs Share Beta using Price Drop.

Table 6. Number of Shares in Liquid and Illiquid Class per Sector.

Sector	1_Liquid	2_Illiquid	% Illiquid
Consumer, Cyclical	42	89	68%
Technology	28	37	57%
Financial	36	135	79%
Basic Materials	11	48	81%
Consumer, Non-cyclical	104	62	37%
Industrial	48	99	67%
Communications	28	32	53%
Utilities	35	11	24%
Energy	8	29	78%
Total	340	542	61%

Table 7. Average of Log (Market Cap) per Class Label and the Sector.

Sector	1_Liquid	2_Illiquid	Gap	Gap%
Consumer, Cyclical	0.84	0.82	0.02	2%
Technology	0.99	0.83	0.16	16%
Financial	1.04	1.13	−0.09	−9%
Basic Materials	0.94	0.86	0.08	8%
Consumer, Non-cyclical	1.15	0.66	0.49	42%
Industrial	0.98	0.73	0.25	26%
Communications	1.41	0.87	0.54	38%
Utilities	1.19	1.14	0.05	4%
Energy	2.02	1.32	0.70	35%
All Shares	1.10	0.90	0.20	18%

Table 8. Average of Share Beta per Class Label and the Sector.

Sector	1_Liquid	2_Illiquid	Gap	Gap%
Consumer, Cyclical	0.95	1.20	0.25	−26%
Technology	0.87	1.05	-0.18	−21%
Financial	1.14	1.32	-0.18	−16%
Basic Materials	0.95	1.09	-0.14	−15%
Consumer, Non-cyclical	0.66	0.79	-0.13	−19%
Industrial	0.93	1.09	-0.16	−17%
Communications	0.83	0.94	-0.12	−14%
Utilities	0.63	0.74	−0.11	−18%
Energy	0.84	0.91	−0.07	−8%
All Shares	0.83	1.10	−0.28	−34%

Table 9. Comparison of Models[5].

Model Name	Acc	Err	Sens.	Spec.	BA	WBA_L	WBA_ILL	AUC
Optimizable Ensemble -Bagged-V6	74.0%	26.0%	57.6%	84.3%	71.0%	64.3%	73.4%	78.0%
Ensemble-Boosted Trees-V6	73.6%	26.4%	60.3%	81.9%	71.1%	65.7%	71.2%	77.0%
Ensemble-Bagged Trees-V6	74.3%	25.7%	60.3%	83.0%	71.7%	66.0%	72.5%	78.0%
Ensemble-RUSBoost-V6	73.2%	26.8%	70.0%	75.3%	72.6%	71.3%	66.8%	78.0%
Ensemble-RUSBoost -V3- PCA 99%	68.5%	31.5%	62.9%	72.0%	67.4%	65.2%	61.8%	74.0%
Ensemble-RUSBoost-V4	73.8%	26.2%	71.8%	75.1%	73.4%	72.6%	67.1%	78.0%
Optimisable Ensemble-V4	75.1%	24.9%	58.2%	85.6%	71.9%	65.1%	75.2%	79.0%
Optimisable Tree	73.2%	26.8%	55.9%	84.1%	70.0%	62.9%	72.7%	74.0%
Logistic Regression	72.6%	27.4%	57.6%	81.9%	69.8%	63.7%	70.5%	77.0%
Optimisable Naïve Bayes	73.8%	26.2%	60.3%	82.3%	71.3%	65.8%	71.7%	78.0%
Optimisable SVM	74.4%	25.6%	55.9%	86.0%	70.9%	63.4%	75.1%	79.0%
Neural Network (Narrow)	70.2%	29.8%	60.6%	76.2%	68.4%	64.5%	65.2%	73.0%

Table **8** shows the average share beta across sector and liquidity classes. For all sectors, a higher beta suggests a more illiquid classification. This makes intuitive sense since higher beta means that when there is market-wide stress, a specific share will have more variance than the market. The consumer, non-cyclical and utilities sectors have the lowest average share beta across all sectors.

Table **9** presents the performance of 12 models against each evaluation metric. Each optimisable model ran 30 iterations of different algorithms and hyperparameters, and the results for the 12 best-performing models are reported. K-fold cross-validation (where K=5) was used for all classification models to prevent overfitting. Without cross-validation, in-sample accuracy would be very high, but performance for out-of-sample predictions would suffer.

When using accuracy (or inversely error) or weighted balance illiquid as a measure, 'Optimisable Ensemble–V4' shows the highest predictive power of all models. However, when the focus is moved to the prediction of each class label, it performs poorly for liquid class, where sensitivity is only 58.2%. Using the preferred measure of a high balanced accuracy and a smaller gap between sensitivity and specificity, 'Ensemble–RUSBoost–V4' becomes the best-performing model, with both classes being correctly predicted more than 70% of the time.

The Confusion Matrix (Fig. **2**) for the selected model (Ensemble–RUSBoost) shows a true liquid class rate (sensitivity) of 71.8% and 75.1% for the true illiquid class rate (specificity). This high prediction power supports the fact that sector, industry, market capitalisation and share beta provide useful information about share liquidity behaviour under conditions of market liquidity stress.

[5] Where a model has 'V6' next to its name, all six features in Table 1.3 were used. Based on several iterations of the model, if prediction power was not much impacted by 90-day trading volume and median market cap percentage dropped, then the model name is listed as either just the model name or with 'V4' added, indicating only the first four features were used in the model.

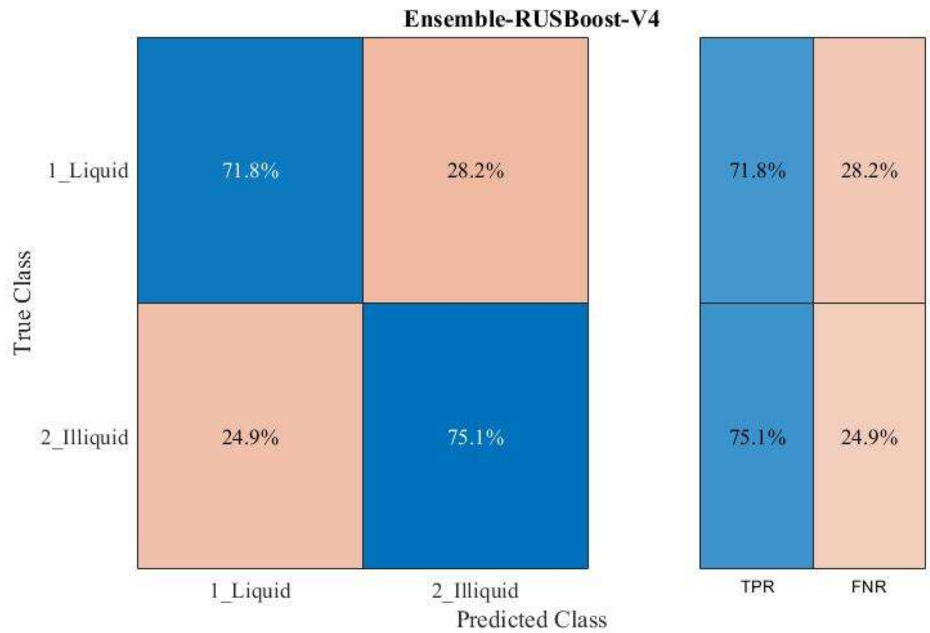


Fig. (2). Confusion Matrix- Ensemble Model RUSBoost.

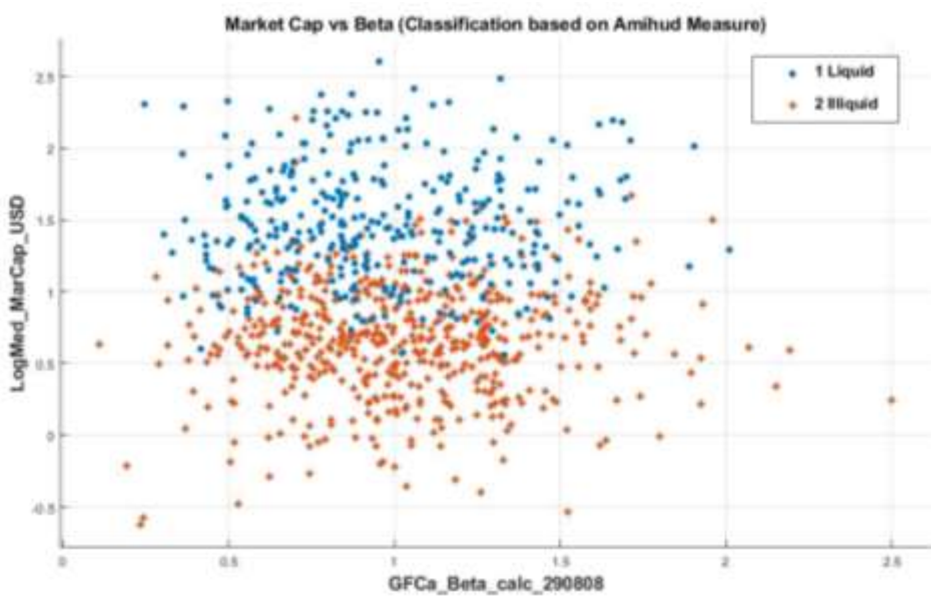


Fig. (3). Original Data Set Market Cap vs Share Beta using Amihud Measure.

6.1. Results Using Amihud Measure for Model Training

To get an initial perspective on the results from the application of the Amihud measure for model training, a scatter plot was produced for market cap and share beta (Fig. **3**). Blue dots represent shares labelled as liquid. High market cap stocks show more liquidity across different beta calculations. When market cap reduces, classification first becomes mixed and then approaches the bottom of the graph as it becomes illiquid. Since the Amihud measure is price impact per USD value traded, for big market size stocks, this impact may be expected to be lower.

When using the Amihud measure, the industrial, basic materials and technology sectors have the highest percentage of illiquid shares. The average log (market cap) for shares classified as liquid is even higher compared to the illiquid class using the Amihud measure for all sectors. The average share beta overall is smaller for the liquid shares group, but the financial, basic materials and energy sectors show the opposite of this. Tables showing details of these results can be found in the Appendix section.

Table **10** presents the performance of 12 models against each evaluation metric when the Amihud measure was used to split shares into class labels. Overall, more models performed well in estimating classification compared to the results in the CMPD. When using accuracy (or inversely error) or weighted balance illiquid as a measure, 'Optimisable Ensemble–V6' shows the highest predictive power. If Optimis-

Table 10. Comparison of Models.

Model Name	Acc	Err	Sens.	Spec.	BA	WBA_L	WBA_ILL	AUC
Optimizable Ensemble -Bagged-V6	85.3%	14.7%	78.8%	89.3%	84.0%	81.4%	83.9%	93.0%
Ensemble-Boosted Trees-V6	83.4%	16.6%	78.5%	86.6%	82.5%	80.5%	80.5%	91.0%
Ensemble-Bagged Trees-V6	84.7%	15.3%	78.8%	88.4%	83.6%	81.2%	82.8%	93.0%
Ensemble-RUSBoost-V6	82.9%	17.1%	80.5%	84.3%	82.4%	81.5%	78.3%	91.0%
Ensemble-RUSBoost -V3- PCA 99%	70.0%	30.0%	70.2%	69.8%	70.0%	70.1%	61.9%	78.0%
Ensemble-RUSBoost-V4	82.7%	17.3%	79.6%	84.5%	82.1%	80.9%	78.3%	90.0%
Optimisable Ensemble-V4	84.1%	15.9%	80.5%	86.4%	83.5%	82.0%	80.6%	92.0%
Optimisable Tree	83.6%	16.4%	79.9%	85.8%	82.9%	81.4%	79.9%	84.0%
Logistic Regression	84.0%	16.0%	78.5%	87.5%	83.0%	80.7%	81.6%	88.0%
Optimisable Naïve Bayes	83.0%	17.0%	73.2%	89.1%	81.1%	77.2%	82.9%	91.0%
Optimisable SVM	85.6%	14.4%	80.2%	89.0%	84.6%	82.4%	83.7%	93.0%
Neural Network (Trilayered)	81.6%	18.4%	75.5%	85.5%	80.5%	78.0%	78.7%	84.0%

able Ensemble is run with four variables instead of six (‘Optimisable Ensemble–V4’), accuracy suffers only very slightly, therefore fewer variables with less computing and data usage would be preferable. Models with four variables all performed reasonably well, except Naïve Bayes, which reported a lower sensitivity measure. Therefore, it can be concluded that several supervised machine learning models produced a high prediction power (above 80%) for liquidity classification using the Amihud Measure.

Based on the above empirical results from models trained using the CMPD or the Amihud measure, it can be concluded that the predictive performance of the ensemble model with RUSBoost algorithm using four features/variables is satisfactory for employment in the liquidity classification problems. Additionally, producing high prediction from these measures supports the fact that under stress, liquidity behaviour of a share is impacted by the sector, industry, market capitalisation and the share beta. These can be used to support liquidity classification, which would help to measure risk sensitivities at the more granular level. Further work can be done using a wide set of liquidity measures and different share features to train models.

7. CONCLUSIONS

This paper employed supervised machine learning models to predict the liquidity classification of common equity shares. Eight hundred and eighty-two unique shares and market data from January 2007 to June 2019 were used in the analysis. The 2007–2008 global financial crisis period following the Lehman Brothers collapse was identified as the most stressful period, and market liquidity measures and features in this period were examined closely to provide further insight into share liquidity behaviour in a market stress environment.

This study showed that the ensemble method with a random undersampling algorithm (Ensemble – RUSBoost) performed comparatively better using preferred metrics such as balanced accuracy and a smaller gap between sensitivity and specificity evaluation metrics. Although this model performed consistently under two liquidity measures used as a response variable, applying the Amihud measure with other supervised machine learning models also showed a high predictive power.

The methodology employed, including transforming liquidity measures to create a classification problem, distinguishes this study from existing literature. It has been shown that supervised machine learning models can be a very useful tool for banks and regulators to further investigate assumptions and initial rules set by the Basel Committee. Another important contribution made is using sector, share beta, industry, and market capitalisation information to predict share liquidity behaviour under stress. Lower beta and higher market cap stocks show more liquid behaviour, and some sectors are more volatile and less liquid than others under market stress.

The model and framework proposed in this study can be applied by financial institutions or regulators to achieve a more granular analysis supported by actual data. This will enable risk sensitivities to be more accurately distinguished, providing the right pricing and funding framework for assets acquired.

Although this study addresses the liquidity classification problem at a more technical level for shares alone, future research can be done to examine other LCR assumptions. Insight could also be gained by integrating the liquidity classification problem into the bank fund transfer pricing mechanism and internal stress testing assumptions.

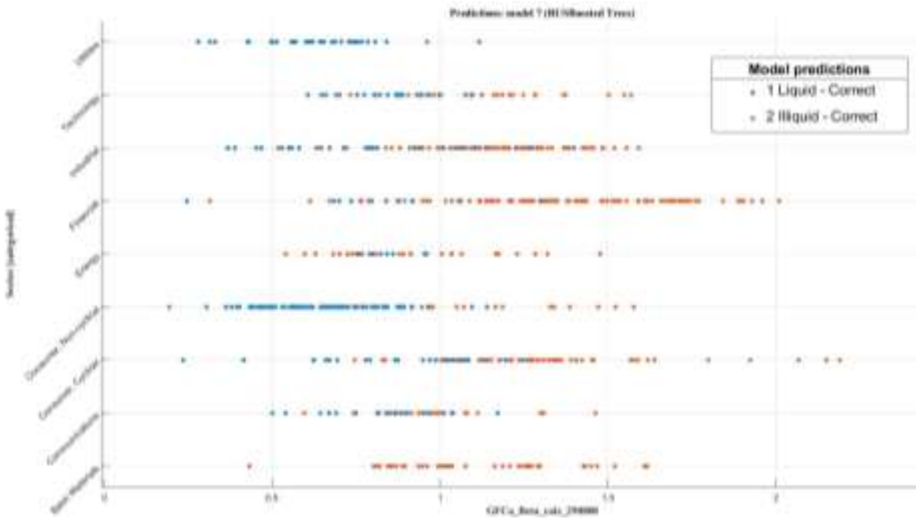
From a policymaking perspective, this study supports the fact that current eligibility criteria in the LCR DA can be further examined, and a more granular approach can be used. This study also shows that machine learning models can be used by regulators to build more granular and risk sensitive assumptions for bank stress testing.

DISCLAIMER

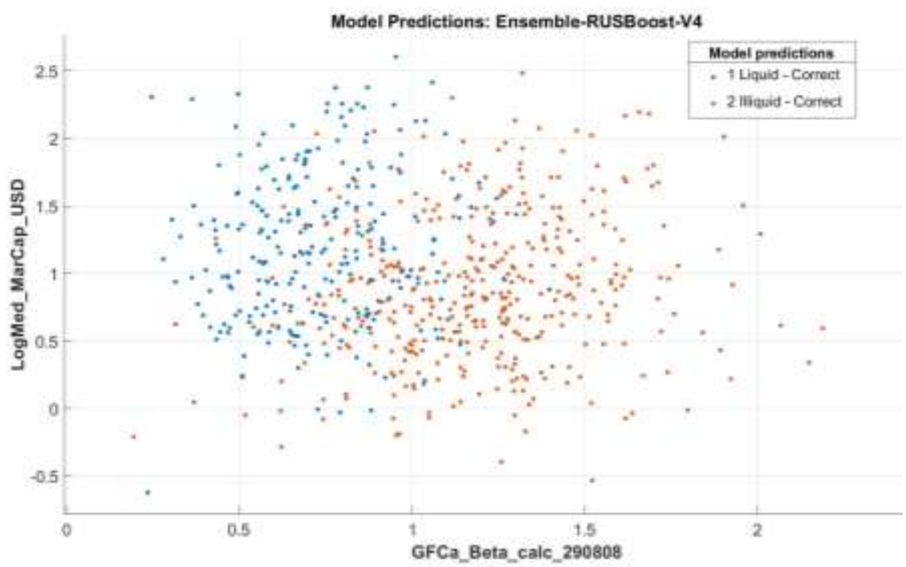
The views and opinions expressed in this paper are those of the authors and they do not necessarily reflect the views of the HSBC Group or Yildiz Technical University.

APPENDIX

Model Predictions Sector vs Beta



Model Predictions Log (Market Cap) vs Beta



Number of Share per Class Label and Sector

Sector	1_Liquid	2_Illiquid	% Illiquid
Consumer, Cyclical	45	86	66%
Technology	21	44	68%
Financial	65	106	62%
Basic Materials	18	41	69%
Consumer, Non-cyclical	74	92	55%
Industrial	42	105	71%
Communications	29	31	52%
Utilities	24	22	48%
Energy	21	16	43%
Total	339	543	62%

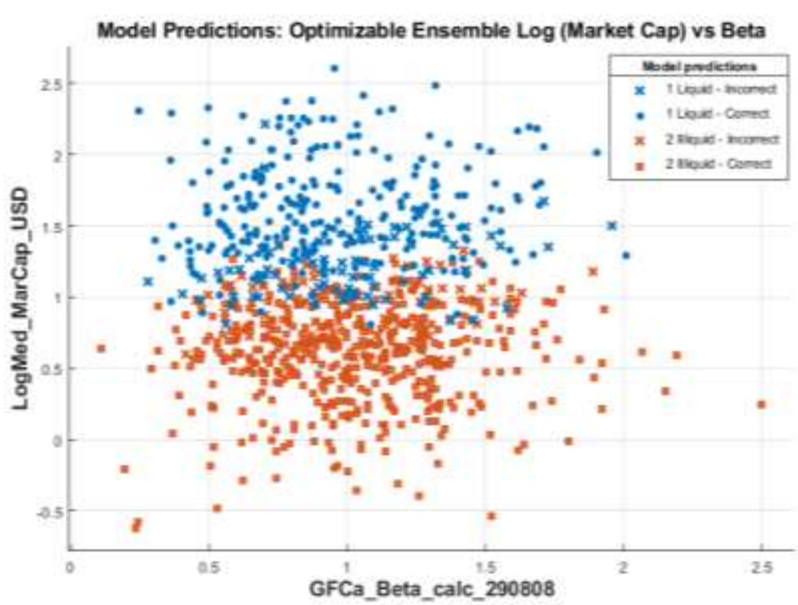
Average of Log (Market Cap) Class Label and Sector

Sector	1_Liquid	2_Illiquid	Gap	Gap%
Consumer, Cyclical	1.28	0.59	0.69	54%
Technology	1.40	0.66	0.75	53%
Financial	1.58	0.82	0.77	48%
Basic Materials	1.39	0.65	0.75	54%
Consumer, Non-cyclical	1.44	0.58	0.86	60%
Industrial	1.39	0.58	0.82	59%
Communications	1.59	0.68	0.91	57%
Utilities	1.43	0.89	0.55	38%
Energy	1.94	0.86	1.07	55%
All Shares	1.48	0.67	0.81	55%

Average of Share Beta per Class Label and Sector

Sector	1_Liquid	2_Illiquid	Gap	Gap%
Consumer, Cyclical	1.07	1.14	-0.07	-7%
Technology	0.96	0.98	-0.02	-2%
Financial	1.31	1.27	0.04	3%
Basic Materials	1.16	1.02	0.14	12%
Consumer, Non-cyclical	0.65	0.75	-0.11	-17%
Industrial	0.99	1.06	-0.06	-6%
Communications	0.88	0.90	-0.02	-2%
Utilities	0.65	0.66	-0.01	-1%
Energy	0.93	0.85	0.08	9%
All Shares	0.96	1.02	-0.06	-7%

Model Prediction (Optimisable Ensemble) Log (Market Cap) vs Share Beta



REFERENCES

Acharya, V. V., & Pedersen, L. H. (2005). Asset Pricing with liquidity risk. *Journal of Financial Economics*, 375-410.

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2017). Measuring Systemic Risk. *The Review fo Financial Studies, Volume 30, Issue 1*, 2-47.

Aldasoro, I., Borio, C., & Drehmann, M. (2018). Early warning indicators fo banking crises; expanding the family. *BIS Quarterly Review*, 29-45.

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

Aras, G., Karaman, Y., & Kazak, E. H. (2020). Efficiency and productivity analysis for intermediary institutions: Turkish capital markets case. *Journal of Capital Markets Study Vol. 4 No.2*, 193-208.

Balaji, A., Ram, D. H., & Nair, B. B. (2018). Applicability of Deep Learning Models for Stock Price Forecasting An Empirical Study on BANKEX Data. *Procedia Computer Science*, 947-953.

Ball, L. M. (2020, November). Liquidity Risk at Large U.S. Banks. 13.

BCBS. (1992). *A Framework for Measuring And Managing Liquidity*.

BCBS. (2000). *Sound Practices for Managing Liquidity in Banking Organisations.*

BCBS. (2008). *Principles for Sound Liquidity Risk Management and Supervision*.

BCBS. (2010, December). *Basel III: A global regulatory framework for more resilient banks and banking systems.* Retrieved from https://www.bis.org/publ/bcbs189_dec2010.pdf

BCBS. (2013). *Basel III: The Liquidity Coverage Ratio and liquidity risk monitoring tools.*

BCBS. (2013). *Evaluating early warning indicators of banking crisis: Satisfying policy requirements.* BIS Working Papers No 421.

BCBS. (2014a, January). *Guidance for Supervisors on Market-Based Indicators of Liquidity.*

BCBS. (2014b). *Basel III: the net stable funding ratio.*

BCBS. (2019, 12 19). Retrieved from https://www.bis.org/basel_framework/chapter/LCR/30.htm?inforce=20191215

Bedowska-Sójka, B. (2019). The dynamics of low-frequency liquidity measures: The developed versus the emerging market. *Journal of Financial Stability 42*, 136-142.

Betz, F., Oprica, S., Peltonen, T. A., & Sarlin, P. (2013, October 11). *Predicting Distress in European Banks.* Retrieved from ECB Working Paper no. 1597: https://ssrn.com/abstract=2338998

BOE. (2019, October). *Machine Learning in UK Financial Services.* Retrieved from https://www.bankofengland.co.uk/-/media/boe/files/report/2019/machine-learning-in-uk-financial-services.pdf

Bonner, C., & Hilbers, P. (2015, January 20). *Global Liquidity Regulation - Why Did it Take so Long?* Retrieved from https://ssrn.com/abstract=2553082

Bordo, M. D. (2010, December). *The Global Financial Crisis of 2007-08: Is it Unprecedented.* Retrieved from https://www.nber.org/system/files/working_papers/w16589/w16589.pdf

Bräuning, M., Malikkidou, D., Scalone, S., & Scricco, G. (2019). *A new approach to Early Warning Systems for small European banks.* ECB Working Papers Series No 2348.

Brunnermeier, M. K., & Pedersen, L. H. (2008). Market Liquidity and Funding Liquidity. *The Review of Financial Studies*, 2201-2238.

Chatzis, S. P., Siakoulis, V., Petropoulos, A., & Evangelos, S. (2018). Forecasting stock market crisis events using deep and statistical. *Expert Systems with Applications*, 353-371.

Demirguc-Kunt, A., & Detragiache, E. (1998). The Determinants of Banking Crises in Developing and Developed Countries. *IMF Staff papers, Volume 45, Number 1*.

EBA. (2013a, February 21). *Discussion Paper on Defining Liquid Assets in the LCR under the draft CRR.*

EBA. (2013b, December 20). *Report on appropriate uniform definitions of extremely high quality assets (extremely HQLA) and high quality liquid assets (HQLA) and on operational requirements for liquid assets under Articl 509 (3) and (5) CRR.* European Banking Authority.

Eross, A., Urquhart, A., & Wolfe, S. (2015, September 29). *An Early Warning Indicator for Liquidity Shortages in the Interbank Market.* Retrieved from Eross, Andrea and Urquhart, Andrew and Wolfe, Simon, An Early Warning Indicator for Liquidity Shortages in the Interbank https://ssrn.com/abstract=2658797

Farag, M., Harland, D., & Nixon, D. (2013, September 17). *Bank of England Quarterly Bullettin.* Retrieved from https://www.bankofengland.co.uk/quarterly-bulletin/2013/q3/bank-capital-and-liquidity

Federal Reserve Bank of St. Louis. (2021, February). Retrieved from https://fred.stlouisfed.org/series/STLFSI2

Fong, K. Y., Holden, C. W., & Trzcinka, C. A. (2017). What are the Best Liquidity Proxies for Global Research? *Review of Finance*, 1355-1401.

FTSE. (2020, September 24). *FTSE Equity Country Classification September 2020*. Retrieved from https://research.ftserussell.com/: https://www.ftserussell.com/equity-country-classification

Gaytán, A., & Johnson, C. A. (2002, October). A Review of the Literature on Early Warning Systems for Banking Crises. *Working Papers Central Bank of Chile 183*. Retrieved from https://ideas.repec.org/p/chb/bcchwp/183.html

Iachini, E., & Nobili, S. (2014, April 23). *An Indicator of Systemic Liquidity Risk in the Italian Financial Markets.* Retrieved from Bank of Italy Occasional Paper No. 217: https://ssrn.com/abstract=2489885

Jones, C. M. (2002, May 23). *A Century of Stock Market Liquidity and Trading Costs.* Retrieved from https://ssrn.com/abstract=313681

Khan, P. Q., Hernes, M., Kuziak, K., Rot, A., & Gryncewicz, W. (2020). Liquidity prediction on Vietnamese stock market using deep learning. *Procedia Computer Science 176*, 2050-2058.

Kliesen, K., & McCracken, M. (2020, March 26). *The St. Louis Fed's Financial Stress Index, Version 2.0*. Retrieved from https://fredblog.stlouisfed.org/2020/03/the-st-louis-feds-financial-stress-index-version-2-0/

Kumar, G., & Misra, A. K. (2015). Closer View at the Stock Market Liquidity: A Literature Review. *Asian Journal of Finance & Accounting*.

Lang, J. H., A., P. T., & Sarlin, P. (2018). *A framework for early-warning modeling with an application to banks.* ECB Working Paper Series.

LCR Delegated Act. (2015). *Official Journal of European Union*, 11.

Leo, M., Sharma, S., & Maddulety, K. (2019). Machine Learning in Banking Risk Management: A Literature Review. *Risks 7,no.1*, 29.

Marshall, B. R., Nguyen, N. H., & Visaltanachoti, N. (2013). Liquidity measurement in frontier markets. *Journal of International Financial Markets, Institutions & Money*, 1-12.

MATLAB. (2019). *Ensemble Algorithms*. Retrieved from https://uk.mathworks.com/help/stats/ensemble-algorithms.html

MATLAB. (2021, 03 09). *Machine Learning in Matlab*. Retrieved from https://uk.mathworks.com/help/stats/machine-learning-in-matlab.html

McKinsey&Company. (2021, 03 09). *An executive's guide to AI*. Retrieved from https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/an-executives-guide-to-ai

Næs, R., Skjeltorp, J. A., & Ødegard, B. A. (2011). Stock Market Liquidity and the Business Cycle. *The Journal of Finance*, 139-176.

Navajas, M. C., & Aaron, T. (2013). *Financial Soundness Indicators and Banking Crisis.* IMF Working paper.

Nosratabadi, S., Mosavi, A., Duan, P., Ghamisi, P., Filip, F., Band, S. S., . . . Gandomi, A. H. (2020). Data Science in Economics: Comprehensive Review of Advanced Machine Learning and Deep Learning Methods. *Mathematics 2020, 8(10)*, 1799.

Odom, M. D., & Sharda, R. (1990). A Neural Network Model for Bankruptcy Prediction. *" International Joint Conference on Neural Networks Volume 2*, (pp. 163-168).

Sarr, A., & Lybek, T. (2002, December). *IMF Working Paper No. 02/232*. Retrieved from Measuring Liquidity in Financial Markets.

scikit-learn. (2020). *1.11. Ensemble methods*. Retrieved from Scikit Learn Documentation: https://scikit-learn.org/stable/modules/ensemble.html

Seiffert, C., Khoshgoftaar, T. M., Hulse, J. V., & Napolitano, A. (2008). RUSBoost: A Hybrid Approach to Alleviating Class Imbalance. *19th International Conference on Pattern Recognition*, (pp. 1-4).

Seiffert, C., Khoshgoftaar, T. M., Van, H. J., & Napolitano, A. (2010). RUSBoost: A Hybrid Approach to Alleviating Class Imbalance. *IEEE Transactions on Systems, Man, and Cybernetics- Part A: Systems and Humans Vol.40 No.1*, 185-197.

Tavana, M., Abtahi Amir-Reza: Di Caprio, D., & Poortarigh, M. (2018). An Artificial Neural Network and Bayesian Network model for liquidity risk assessment in banking. *Neurocomputing*, 2525-2554.

Vayanos, D., & Jiang, W. (2012). Chapter 19- Market Liquidity—Theory and Empirical Evidence. In *Handbook of the Economics and Finance Volume 2, Part B* (pp. 1289-1361).

Venkat, S., & Baird, S. (2016). *Liquidity Risk Management A Practitioner's Perspective.* Wiley Finance Series.

Received: May 02, 2023 Revised: May 10, 2023 Accepted: Sep 04, 2023