Academic Journal of Applied Sciences
ISSN: 3105-0549 (Print), ISSN: 3105-0557 (Online) | Vol. 4, No. 1, 2026

Liquidity Stress Prediction Using Market Microstructure Indicators: Evidence from the Chinese Stock Market

Alessia Gallo

Department of Computer Science, University of Turin, Turin, Italy

Abstract: Market liquidity can evaporate abruptly, yet systematic evidence on whether daily market microstructure indicators anticipate market-wide liquidity stress remains scarce, especially for markets with daily price limits. This paper constructs a liquidity stress index (LSI) from the cross-sectional median Amihud ratio of 5,571 Chinese A-shares over 2005–2026 and predicts stress episodes from 33 market-level indicators that aggregate bid–ask spread estimators, price-impact and trading-activity measures, range-based volatility, and order-flow signatures such as limit-down breadth, one-price locks and sell-side turnover. Range-based spread estimators degenerate on locked limit-down days, whereas the price-impact-based LSI identifies the major stress episodes of the sample. In a walk-forward evaluation over 2012–2026 with quarterly re-estimation and an embargo, a regularized logistic model on microstructure indicators predicts stress within the next five trading days with an AUC of 0.808 and an average precision of 0.663, compared with 0.795 and 0.621 for a heterogeneous autoregressive benchmark, and raises the hit rate at a 10% false alarm rate from 46.0% to 50.5%. The gains persist across horizons from one to twenty days, alternative stress thresholds and stress onsets. Price-impact, trading-activity, volatility and order-flow indicators carry the predictive content, whereas spread estimators add little; sell-side order imbalance and overnight gap risk rise ahead of stress episodes.

Keywords: Market microstructure, liquidity stress, Amihud illiquidity, bid–ask spread estimators, price limits, early-warning system, machine learning.

1. Introduction

Market liquidity—the ability to trade quickly and in size at low cost—is a central determinant of transaction costs, asset prices and financial stability [1–4]. Liquidity is also fragile: it tends to evaporate precisely when it is most needed, because falling prices tighten the funding constraints of liquidity providers and trigger self-reinforcing liquidity spirals [5, 6]. Empirically, market-wide liquidity deteriorates after market declines and during periods of elevated uncertainty [7–9] and co-moves strongly across securities [10–12], so that liquidity shocks constitute a systematic risk [13, 14]. Episodes such as the 2010 flash crash showed how quickly liquidity can disappear in modern electronic markets [15, 16].

Anticipating such episodes is valuable for regulators that calibrate market-stabilization tools, for exchanges and market makers that manage inventory and margin risk, and for investors that must liquidate positions. Composite financial stress indices [17, 18] and early-warning systems [19–22] provide a mature toolkit for monitoring systemic risk, but they are typically built on macro-financial variables observed at monthly or quarterly frequencies. Market microstructure offers a complementary set of daily indicators—bid–ask spread estimators, price-impact ratios, range-based volatility and proxies for one-sided order flow—that can be computed for every listed stock from publicly available daily data [23–26]. Whether, and which, of these indicators carry advance information about market-wide liquidity stress is an open empirical question, particularly in emerging markets where trading frictions are pronounced [27, 28].

The Chinese A-share market is a natural laboratory for this question. It is one of the largest equity markets in the world, its trading is dominated by retail investors, and its daily price limits mechanically truncate price discovery [29–31]. The market has experienced repeated liquidity crises, including the 2015 crash and the subsequent government rescue [32, 33], the January 2016 circuit-breaker episode, and the sell-offs of February 2020 and early 2024. Price limits also create a measurement problem. On days when most stocks are locked at the lower limit, intraday ranges collapse and range-based spread estimators report spreads close to zero although liquidity has vanished. On 3 February 2020, for example, 83.9% of the stocks in our sample closed at the limit-down price and 57.6% traded at a single locked price; the cross-sectional median Parkinson range was exactly zero and the average Corwin–Schultz spread fell to 0.19%, whereas the Amihud-based stress index introduced below reached its highest value of the entire sample. A stress measure for such a market must therefore remain informative under locked limit moves.

This paper studies the prediction of market-wide liquidity stress from daily microstructure indicators. Using daily data on 5,571 Shanghai and Shenzhen A-shares from 2005 to 2026, we compute stock-level spread, price-impact, volatility and order-flow measures, aggregate them into 33 market-level indicators, and define a liquidity stress index (LSI) as the rolling z-score of the market-wide median Amihud ratio, the daily price-impact proxy that best tracks high-frequency benchmarks [34, 35]. We forecast whether a stress day occurs within the next h trading days with regularized logistic regression, random forests, gradient boosting, a multilayer perceptron and a probability-averaging ensemble, and compare them with persistence and heterogeneous autoregressive (HAR) benchmarks [36] in a strictly out-of-sample walk-forward design over 2012–2026.

The contributions are threefold. First, we construct a price-impact-based stress index and a set of market-level indicators that remain informative in a market with daily price limits, including the breadth of limit-down hits and one-price locks,

trading suspensions and the share of turnover executed in declining stocks. Second, we show that microstructure indicators carry predictive information beyond the persistence of illiquidity: at a five-day horizon, the microstructure-augmented logit raises the average precision from 0.621 to 0.663 and the true positive rate at a 10% false positive rate from 46.0% to 50.5% relative to the HAR benchmark, and the gains hold across horizons, stress thresholds and stress onsets. Third, ablation and permutation-importance analyses show that price-impact, trading-activity, volatility and order-flow indicators drive the forecasts, that a model based on volatility and return indicators alone already exceeds the HAR benchmark, and that range-based spread estimators contribute little incremental information in a price-limit market.

The remainder of the paper is organized as follows. Section II reviews related work. Section III describes the data, the indicators and the forecasting design. Section IV reports the empirical results, Section V discusses their implications, and Section VI concludes.

2. Related Work

2.1. Low-Frequency Liquidity Measurement

Because intraday quote and trade data are costly and unavailable over long histories in many markets, a large literature estimates liquidity from daily data. Roll [23] infers the effective spread from the negative serial covariance of price changes, Lesmond *et al.* [37] exploit the frequency of zero returns, Amihud [24] measures price impact as the absolute return per unit of trading value, and Pástor and Stambaugh [13] and Liu [38] propose reversal- and turnover-based measures. Range-based estimators use daily highs and lows: Corwin and Schultz [25] separate the spread and variance components of one- and two-day ranges, Abdi and Ranaldo [26] add closing prices, and Ardia *et al.* [39] derive an efficient estimator from open, high, low and close prices. These estimators are closely related to range-based volatility measures [40–42]. Horse races against intraday benchmarks show that daily price-impact proxies, the Amihud ratio in particular, perform well in the United States and globally [34], [35], and low-frequency measures have been applied to emerging equity markets [27, 28] and to foreign exchange [43].

2.2. Market-Wide Liquidity Dynamics and Liquidity Crises

Liquidity has a strong common component across securities [10–12, 44] that is priced as a systematic risk factor [13, 14, 45, 46]. Aggregate liquidity responds to trading activity, volatility and order imbalances [7, 47, 48], declines after negative market returns [8], and dries up when intermediaries become capital constrained [5, 6, 9, 49]. Volume and volatility are tightly linked [50, 51], and liquidity also reflects investor sentiment and retail order flow [52, 53]. In electronic markets, liquidity crashes can unfold within minutes, as in the 2010 flash crash [15], and order-flow toxicity metrics have been proposed as early indicators [16]. For China, studies document the 2015 crash and the state intervention that followed [32, 33], the distortions created by daily price limits [29, 31], and the evolution of the market's microstructure [54].

2.3. Financial Stress Indices and Early-Warning Models

Financial stress indices aggregate standardized market indicators into a single measure of contemporaneous stress, as in the CISS of Holló *et al.* [17], and are used to date stress episodes [18]. Early-warning models predict such events from leading indicators, ranging from the signals approach [19] to logit models and machine-learning classifiers [20–22]. The evidence on machine learning is mixed: ensembles improve robustness [20], whereas well-specified logit models are hard to beat out of sample [21]. In asset pricing, tree ensembles and neural networks capture nonlinear interactions among liquidity and volatility signals [55], and financial applications call for purged and embargoed validation to avoid look-ahead bias [56]. Our study connects these strands by forecasting market-wide liquidity stress at a daily frequency from microstructure indicators under the evaluation discipline of the early-warning literature.

3. Data and Methodology

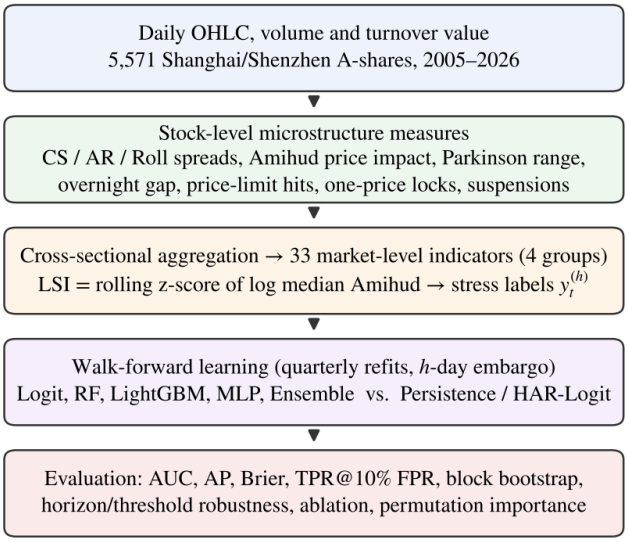


Figure 1. Research framework for liquidity stress prediction from market microstructure indicators.

3.1. Data and Sample

We use daily open, high, low and close prices adjusted for corporate actions, trading volume and trading value for all A-shares listed on the Shanghai and Shenzhen stock exchanges, obtained from an open-source Qlib-format database. The sample runs from 4 January 2005 to 13 November 2026 (5,309 trading days) and covers 5,571 stocks, including delisted firms; stocks listed on the Beijing Stock Exchange are excluded. On each day we retain listed stocks with valid prices and positive turnover, drop the first 60 trading days after each initial public offering to remove post-listing limit-up sequences, and discard daily returns whose absolute value exceeds 22%, which no price-limit regime permits. The resulting universe contains 2,952 stocks per day on average, rising from about 1,100 in 2005–2006 to more than 5,000 in 2024–2026. Fig. 1 summarizes the research framework.

3.2. Stock-Level Microstructure Measures

Let $O_{i,t}$, $H_{i,t}$, $L_{i,t}$ and $C_{i,t}$ denote the open, high, low and close prices of stock i on day t, $r_{i,t}$ its return and $V_{i,t}$ its trading value. All measures use only information available at the close of day t. The Corwin–Schultz spread [25] is computed over the

two-day window $\{t-1, t\}$ after adjusting day-t prices for overnight gaps. With β the sum of the two squared daily log ranges and γ the squared log range of the two-day high and low (stock and day subscripts omitted),

$$\alpha = \frac{\sqrt{2\beta} - \sqrt{\beta}}{3 - 2\sqrt{2}} - \sqrt{\frac{\gamma}{3 - 2\sqrt{2}}}, \quad S^{CS} = \frac{2(e^{\alpha} - 1)}{1 + e^{\alpha}} \tag{1}$$

And negative estimates are set to zero. The Abdi–Ranaldo estimator [26] uses the log close c_t and the log mid-range $\eta_t = (\ln H_t + \ln L_t)/2$:

$$S_t^{AR} = \sqrt{\max\{4(c_{t-1} - \eta_{t-1})(c_{t-1} - \eta_t),\ 0\}} \tag{2}$$

The Roll [23] spread is estimated from log price changes Δp over a rolling 21-day window,

$$S_t^{R} = 2\sqrt{\max\{-\mathrm{Cov}(\Delta p_{\tau},\ \Delta p_{\tau-1}),\ 0\}} \tag{3}$$

Price impact is measured by the Amihud [24] ratio with V expressed in units of 10^8 CNY, and volatility by the Parkinson [40] range:

$$\mathrm{ILLIQ}_{i,t} = \frac{|r_{i,t}|}{V_{i,t}}, \qquad \sigma_{i,t}^{P} = \frac{\ln(H_{i,t}/L_{i,t})}{\sqrt{4\ln 2}} \tag{4}$$

We also record the absolute overnight gap $|\ln(O_{i,t}/C_{i,t-1})|$, price-limit hits, one-price locks and trading suspensions. A stock hits its lower (upper) limit when $r_{i,t} \le -(\ell_{i,t} - 0.5\%)$ ($r_{i,t} \ge \ell_{i,t} - 0.5\%$), where the limit $\ell_{i,t}$ is 10% for main-board stocks and 20% for STAR Market stocks and for ChiNext stocks after 24 August 2020. A one-price lock is a day with $(H_{i,t} - L_{i,t})/C_{i,t} < 10^{-4}$ and $|r_{i,t}| \ge 4.5\%$, which also covers the 5% limit of special-treatment stocks. A listed stock without valid trading data is classified as suspended.

3.3. Market-Level Indicators

The stock-level measures are aggregated across the daily universe. The zero-floored spread estimators are averaged cross-sectionally; the skewed Amihud ratio, Parkinson range and overnight gap are summarized by cross-sectional medians; dispersion is measured by interquartile ranges; and breadth indicators are the shares of stocks with limit-down or limit-up hits, one-price down- or up-locks, negative returns and suspensions. We further compute the logarithm of aggregate trading value, the share of trading value executed in declining stocks as a daily proxy for sell-side order imbalance [47], the equal-weighted market return, its realized volatility and the drawdown from the 60-day high. Level variables with secular trends are transformed into rolling 250-day z-scores, so that each indicator is measured relative to its recent history. Table I lists the resulting 33 predictors in four groups.

Table I. Market-Level Microstructure Indicators

Group (No.)	Indicators
Spread (6)	Cross-sectional mean CS and AR spreads (daily and 5-day mean), 21-day Roll spread, interquartile range of CS spreads
Price impact and activity (7)	LSI and its 5- and 22-day means; interquartile range of log Amihud ratios; log aggregate trading value, its deviation from the 20-day mean and its 5-day change
Volatility and return (10)	Median Parkinson range (daily, 5-day mean), median overnight gap, 5- and 22-day realized volatility, 1-, 5- and 20-day market returns, 60-day drawdown, return dispersion
Order flow and frictions (10)	Shares of limit-down hits (daily, 5-day mean), limit-up hits, down-locks, up-locks, decliners (daily, 5-day mean), sell-side turnover share (daily, 5-day mean), 5-day change in suspension share

3.4. Liquidity Stress Index and Stress Events

We measure market-wide liquidity stress through the price-impact dimension. Let x_t be the logarithm of the cross-sectional median Amihud ratio on day t. The liquidity stress index is

$$\mathrm{LSI}_t = \frac{x_t - \mu_t}{s_t} \tag{5}$$

Where μ_t and s_t are the mean and standard deviation of x over the trailing 250 trading days. The Amihud ratio is the best daily cost-per-volume proxy in global horse races [35] and, unlike range-based estimators, it rises sharply on locked limit-down days, when large absolute returns coincide with collapsing trading value. A day is a stress day when LSI_t exceeds its 90th percentile $\tau = 1.41$, which yields 511 stress days among 5,110 days with a valid index. The prediction target for horizon h is

$$y_t^{(h)} = \mathbf{1}\{\max_{1 \le k \le h} \mathrm{LSI}_{t+k} > \tau\} \tag{6}$$

i.e., an indicator of at least one stress day within the next h trading days. The main horizon is $h = 5$ (one trading week), for which the event frequency in the test period is 28.6%.

3.5. Forecasting Models

Two benchmarks use only the history of the stress index: a persistence logit with LSI_t as its single predictor, and a HAR-logit with the daily, weekly (5-day) and monthly (22-day) averages of the LSI, which mirrors the heterogeneous autoregressive structure of volatility forecasting [36]. Against these benchmarks we estimate four classifiers on all 33 indicators: an L2-regularized logistic regression (Logit, inverse regularization strength 0.1) on standardized inputs; a random forest (RF) with 300 trees and at least 10 observations per leaf [57]; a gradient-boosted tree model (LightGBM) with 400 trees of at most seven leaves, learning rate 0.02 and row and column subsampling [58], [59]; and a multilayer perceptron (MLP) with hidden layers of 32 and 16 units, an L2 penalty and early stopping. Motivated by the robustness of model averaging in early-warning applications [20], the Ensemble averages the predicted probabilities of Logit, RF and LightGBM:

$$p_t^{\mathrm{Ens}} = \frac{1}{3}\left(p_t^{\mathrm{Logit}} + p_t^{\mathrm{RF}} + p_t^{\mathrm{LGBM}}\right) \tag{7}$$

All models are implemented in Python with scikit-learn [60] and LightGBM. Hyperparameters are set to conservative values in advance and are not optimized on the test sample.

3.6. Evaluation Protocol

Forecasts are produced with an expanding-window walk-forward scheme. The first test day is 4 January 2012, and all

models are re-estimated every 63 trading days (one quarter) on all data available at that point. Because the label of day t depends on the index up to $t + h$, training observations whose label window overlaps the test block are removed (an h-day embargo [56]), so that no future information enters estimation. The test period ends in November 2026 and contains 3,603 daily forecasts at $h = 5$. We report the area under the ROC curve (AUC) [61], the average precision (AP), which is more informative under class imbalance [62], the Brier score, and the true positive rate at a 10% false positive rate (TPR@10%FPR), a natural operating point for an alarm system. AUC differences are assessed with a moving-block bootstrap with 20-day blocks and 1,000 replications [63], which preserves the serial dependence of forecasts and labels. Indicator relevance is measured by permutation importance [57], the decrease in test AUC when one indicator is randomly permuted across test days.

4. Empirical Results

4.1. Stress Episodes and Indicator Dynamics

Fig. 2 plots the LSI together with the 511 stress days. Stress days cluster in well-known episodes: the global financial crisis of 2008, the crash of mid-2015 and its aftershock in August 2015, the circuit-breaker turmoil of January 2016, the trade-tension sell-offs of 2018, the reopening of the market after the Lunar New Year holiday on 3 February 2020, the sell-off of April 2022 and the small-capitalization liquidity crunch of January–February 2024. The index is persistent but far from deterministic (first-order autocorrelation 0.62), and the stress days form 88 distinct episodes when separated by gaps of more than two weeks, which provides a rich set of events for out-of-sample evaluation.

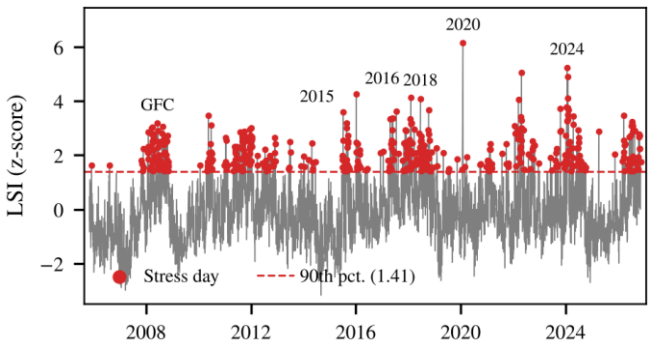


Figure 2. Liquidity stress index (LSI) of the A-share market. Red dots mark stress days (LSI above its 90th percentile, dashed line). Labels mark the global financial crisis (GFC), the 2015 crash, the January 2016 circuit-breaker episode, the 2018 sell-offs, the February 2020 reopening and the early-2024 small-cap crunch.

Table II reports average indicator levels on normal days, in the five days preceding a stress onset (pre-stress) and on stress days. During stress, the log median Amihud ratio rises by about one unit (from −3.85 to −2.75), range volatility increases by roughly one half, the shares of decliners, limit-down hits and down-locks surge, and aggregate trading value falls. The pre-stress column reveals the leading signals. In the week before a stress onset, the share of turnover executed in declining stocks rises from 41.1% to 49.9%, the share of decliners from 44.0% to 49.1%, the median overnight gap from 0.46% to 0.60% and the limit-down share from 0.50% to 0.68%. The spread estimators, by contrast, barely move before stress: the CS spread increases from 0.96% to 1.00% and the Roll spread is unchanged, consistent with their limited ability to register one-sided, limit-constrained order flow.

Table II. Microstructure Indicators by Market State

Indicator	Normal	Pre-stress	Stress
CS spread (%)	0.96	1.00	1.01
AR spread (%)	0.96	0.98	1.24
Roll spread (%)	1.13	1.12	1.13
log Amihud	-3.85	-3.72	-2.75
Parkinson vol. (%)	2.11	2.17	3.13
Overnight gap (%)	0.46	0.60	0.93
log turnover value	27.90	27.86	27.76
Limit-down share (%)	0.50	0.68	5.85
Down-lock share (%)	0.08	0.14	0.36
Decliner share (%)	44.04	49.12	66.12
Sell-side turnover share (%)	41.06	49.94	67.89
Suspension share (%)	4.72	4.13	4.82
Days (N)	3551	1048	511

Pre-stress: the five trading days before a stress onset that are not themselves stress days. Spreads, volatility, gaps and shares in percent; trading value in log CNY.

4.2. Out-of-Sample Prediction Performance

Table III reports the out-of-sample results at $h = 5$. Both benchmarks are informative—the persistence logit attains an AUC of 0.756 and the HAR-logit 0.795—which confirms that liquidity stress is persistent. Adding the microstructure indicators improves every metric. The regularized logit attains the highest AUC (0.808), AP (0.663) and TPR@10%FPR (0.505) and the lowest Brier score (0.151); the Ensemble is close behind (AUC 0.805, AP 0.662), followed by the RF (AUC 0.803). Relative to the HAR benchmark, the AP of the logit increases by 0.042, a relative gain of 6.7%. The moving-block bootstrap places the AUC gain of the logit at 0.014 with a 95% interval of [−0.002, +0.029] and a one-sided p-value of 0.043, and all models except the MLP improve on the persistence benchmark by 0.03–0.05 in AUC ($p < 0.01$).

Table III. Out-of-Sample Prediction of Liquidity Stress (h = 5)

Model	AUC	AP	Brier	TPR	ΔAUC	p
Persistence	0.756	0.561	0.172	0.363	–	–
HAR-Logit	0.795	0.621	0.158	0.460	–	–
Logit	0.808	0.663	0.151	0.505	+0.014	0.043
RF	0.803	0.656	0.152	0.486	+0.008	0.114
LightGBM	0.788	0.638	0.160	0.472	−0.006	0.730
MLP	0.777	0.629	0.164	0.468	−0.019	0.929
Ensemble	0.805	0.662	0.151	0.501	+0.011	0.081

Test period January 2012–November 2026 (3603 forecasts, event frequency 28.6%). TPR: true positive rate at a 10% false positive rate. ΔAUC: AUC difference to the HAR-logit; p: one-sided moving-block bootstrap p-value.

Fig. 3 shows that the advantage of the microstructure-based models is concentrated in the region of low false positive rates and high precision that matters for an alarm system. The nonlinear learners do not dominate the linear model: LightGBM and the MLP trail the logit, in line with early-warning evidence that parsimonious logit models are hard to beat out of sample [21]. With between about 1,500 and 5,000 daily training observations per refit and strongly correlated predictors, the bias–variance trade-off favours the regularized

linear specification, while the Ensemble retains most of its accuracy and adds robustness across horizons.

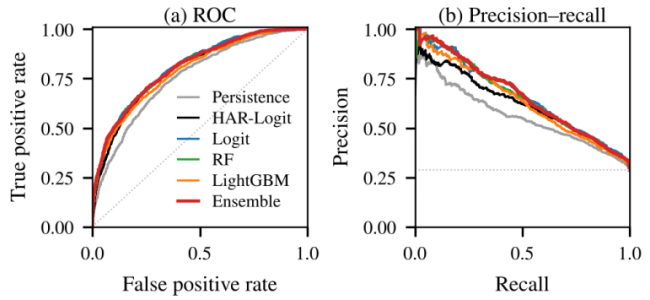


Figure 3. Out-of-sample ROC (a) and precision–recall (b) curves for stress prediction at h = 5 over 2012–2026. The dotted line in (b) marks the event frequency.

4.3. Horizons, Stress Thresholds and Stress Onsets

Fig. 4 reports AUC and AP for horizons of 1, 5, 10 and 20 trading days. AP rises with the horizon because the event frequency increases from 9.6% at $h = 1$ to 55.1% at $h = 20$, whereas AUC declines gradually as the forecasting problem becomes harder. The microstructure-based models outperform the HAR benchmark at every horizon. At $h = 1$ the Ensemble attains the highest AUC (0.836 versus 0.812) and AP (0.384 versus 0.323), and at $h = 20$ the logit remains ahead (AUC 0.770 versus 0.761; AP 0.810 versus 0.799). Tree ensembles are most useful at short horizons, where the RF reaches an AUC of 0.833, whereas the regularized logit is the most stable model at longer horizons.

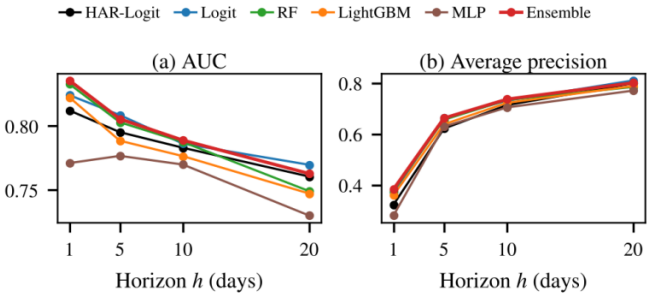


Figure 4. Out-of-sample AUC (a) and average precision (b) by forecast horizon h.

Table IV varies the stress threshold and evaluates stress onsets. With an 85th-percentile threshold the event frequency rises to 39.2% and the logit improves the AUC of the HAR benchmark from 0.798 to 0.813. With a 95th-percentile threshold, which isolates the most severe episodes, the gain widens to 0.020 in AUC (0.829 versus 0.809) and 0.045 in AP. The onset analysis restricts the evaluation to days that are not themselves stress days, so that the forecast must anticipate a transition into stress rather than its continuation. Here the persistence benchmark loses much of its power (AUC 0.716), whereas the logit and the Ensemble attain AUCs of 0.782 and 0.778 and APs of 0.556 and 0.560, against 0.765 and 0.508 for the HAR benchmark. Microstructure indicators are therefore most valuable where early warning matters, namely before the index itself signals stress.

Table IV. Robustness to Stress Threshold and Stress Onsets (h = 5)

Setting	Metric	Persist.	HAR	Logit	LGBM
85th pct. (39.2%)	AUC	0.758	0.798	0.813	0.796
	AP	0.674	0.732	0.756	0.740
90th pct. (28.6%)	AUC	0.756	0.795	0.808	0.788
	AP	0.561	0.621	0.663	0.638
95th pct. (16.6%)	AUC	0.781	0.809	0.829	0.818
	AP	0.437	0.500	0.545	0.484
Onsets (24.3%)	AUC	0.716	0.765	0.782	0.760
	AP	0.408	0.508	0.556	0.536

Percentiles refer to the stress threshold τ; event frequencies in parentheses. Onsets: test days on which the LSI is below the 90th-percentile threshold.

4.4. Which Microstructure Indicators Matter?

Fig. 5 reports the permutation importance of the 12 most relevant indicators for the Ensemble at $h = 5$, and Table V reports feature-group ablations for the logit and LightGBM. The smoothed stress index (its 5- and 22-day means) ranks first, reflecting the persistence of illiquidity, followed by aggregate trading value, the 60-day drawdown, the 20-day market return, the 5-day average limit-down share, return dispersion and short-term realized volatility. The ablations confirm this ranking at the group level. Removing the price-impact and activity group causes the largest loss (logit AUC 0.789, AP 0.621), removing the volatility and return group the second largest (AUC 0.802, AP 0.650), and removing the order-flow group lowers the AP to 0.656. Removing the spread estimators leaves the logit unchanged (AUC 0.808, AP 0.666), and spread estimators alone are barely informative (AUC 0.549).

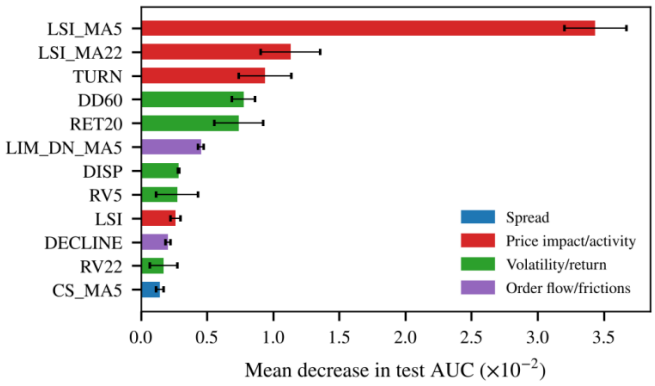


Figure 5. Permutation importance of the 12 most relevant indicators for the Ensemble at h = 5 (mean decrease in test AUC over three permutations; error bars show one standard deviation).

Two further results stand out. A logit that uses only the ten volatility and return indicators (AUC 0.799, AP 0.646) already exceeds the HAR benchmark, and a logit that excludes the entire LSI history (AUC 0.785, AP 0.625) performs on par with it. Market-level microstructure signals therefore contain much of the information in the stress index's own past and add to it when combined. The same pattern holds for LightGBM, for which order-flow indicators alone reach an AUC of 0.692 compared with 0.594 for spread estimators alone.

Table V. Feature-Group Ablation (h = 5)

Specification	Logit AUC	Logit AP	LGBM AUC	LGBM AP
All 33 indicators	0.808	0.663	0.788	0.638
Without spread	0.808	0.666	0.786	0.632
Without price impact and activity	0.789	0.621	0.778	0.600
Without volatility and return	0.802	0.650	0.789	0.639
Without order flow and frictions	0.806	0.656	0.792	0.643
Without LSI history	0.785	0.625	0.775	0.619
Spread only	0.549	0.342	0.594	0.384
Price impact and activity only	0.790	0.616	0.778	0.605
Volatility and return only	0.799	0.646	0.777	0.599
Order flow and frictions only	0.666	0.466	0.692	0.484

LSI history: the LSI and its 5- and 22-day means.

4.5. Case Studies

Fig. 6 traces the five-day-ahead stress probabilities of the Ensemble and the HAR benchmark around two major episodes. In 2015, both probabilities rise sharply when the June sell-off turns into a crash, remain elevated through the August turmoil and the circuit-breaker episode of January 2016, and decline quickly in the calm fourth quarter of 2015, which limits false alarms. In the 2024 episode, the Ensemble probability stays above the HAR probability for most of December 2023 and early January 2024, as small-capitalization stocks deteriorated gradually ahead of the late-January stress cluster. Across all stress onsets in the test period, the Ensemble assigns an average probability of 0.418 to the five days preceding an onset, compared with 0.378 for the HAR benchmark and 0.309 for the persistence model, while its average probability on other non-stress days is only marginally higher (0.189 versus 0.177).

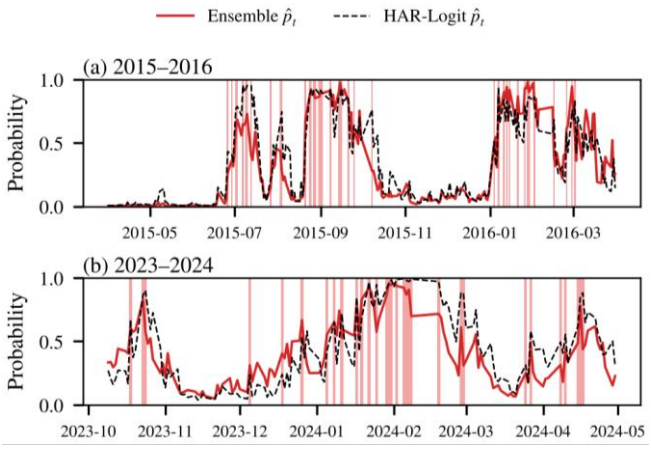


Figure 6. Five-day-ahead stress probabilities of the Ensemble and the HAR-logit around the 2015–2016 (a) and 2023–2024 (b) episodes. Shaded bars mark realized stress days.

5. Discussion

The results carry three implications. First, in a market with daily price limits, liquidity stress should be measured through price impact rather than through range-based spreads. Price limits censor intraday ranges precisely when selling pressure is strongest, so that spread estimators understate illiquidity on the most stressful days, whereas the Amihud ratio captures the combination of large price moves and collapsing turnover. Second, the leading information resides in order flow and trading activity. A rising share of turnover in declining stocks, wider overnight gaps, more frequent limit-down hits and falling trading value precede stress, consistent with theories in which one-sided order flow exhausts the risk-bearing capacity of liquidity providers [5], [47] and with evidence that liquidity deteriorates after market declines [8]. Third, a regularized logit on economically motivated indicators is competitive with nonlinear learners, and averaging it with tree ensembles yields a robust alarm whose accuracy is close to that of the best single model at each horizon.

Because all indicators are computed from end-of-day data that exchanges disseminate publicly, the framework can be updated daily at negligible cost and integrated into the surveillance toolkits of exchanges and regulators. A five-day stress probability can support the timing of margin adjustments, the preparation of liquidity-support facilities and the calibration of market-stabilization measures before stress materializes, and it can serve market makers and institutional investors as an input to inventory and execution risk management.

6. Conclusion

This paper shows that daily market microstructure indicators predict market-wide liquidity stress in the Chinese A-share market. A price-impact-based stress index identifies the major liquidity crises of 2005–2026, and models that combine price-impact, trading-activity, volatility and order-flow indicators forecast stress up to four weeks ahead with AUCs between 0.770 and 0.836, improving on autoregressive benchmarks at every horizon, for alternative stress thresholds and for stress onsets. Sell-side order imbalance, overnight gap risk and limit-down breadth emerge as leading signals, whereas range-based spread estimators add little in a market with daily price limits. Future work can extend the framework to intraday order-book data, to spillovers across equity, bond and derivatives markets, and to other markets with price limits.

References

[1] Wei, K., Lu, Z., & Zou, X. (2024). A scalable causal machine learning framework for data-driven decision optimization. AI and Data Science Journal, 1(1), 122–131.

[2] Hu, X., Guo, L., Wang, J., & Liu, Y. (2025). Computational fluid dynamics and machine learning integration for evaluating solar thermal collector efficiency-based parameter analysis. Scientific Reports, 15, 24528.

[3] Wang, J., Tan, Y., Jiang, B., Wu, B., & Liu, W. (2025). Dynamic marketing uplift modeling: A symmetry-preserving framework integrating causal forests with deep reinforcement learning for personalized intervention strategies. Symmetry, 17(4), 610.

[4] Ping, W., & Fan, H. (2023). Beyond isolated ratios: Accrual–revenue–cash flow indicators and interaction-aware learning for detecting financial reporting irregularities in U.S. public companies. World Journal of Economics and Business Research, 1(1), 46–55.

[5] Jiao, Y., & Liang, Y. (2024). An integrated process-mining and explainable machine-learning framework for bottleneck and internal-control risk detection in the multi-entity financial close. Social Science and Management, 1(3), 96–107.

[6] Zhang, F., & Yang, J. (2022). Accuracy-aware model compression for real-time visual perception under resource constraints. Journal of Computer Science and Electrical Engineering, 4(1), 45–52.

[7] Wang, G., Li, P., Feng, E., & Mo, T. (2024). Uncertainty-aware LLM-as-a-judge evaluation for reliable search and recommendation systems: Calibrated confidence and impact-aware selective human review. AI and Data Science Journal, 1(1), 112–121.

[8] Xing, S., Wang, Y., & Liu, W. (2025). Self-adapting CPU scheduling for mixed database workloads via hierarchical deep reinforcement learning. Symmetry, 17(7), 1109.

[9] Bian, M., Wang, Y., & Lin, Y. (2024). Explainable supplier risk prediction for critical telecommunications infrastructure procurement: An empirical XGBoost–SHAP framework. Social Science and Management, 1(1), 64–77.

[10] Wang, J., Liu, J., Ge, Y., & Jiang, B. (2023). Risk-aware ensemble learning for credit default prediction under economic uncertainty. World Journal of Economics and Business Research, 1(1), 18–32.

[11] Wang, G., Li, P., Feng, E., Shen, A., & Rhee, M. (2026). A dual-process memory-augmented large language model agent for explainable sequential recommendation. IEEE Access.

[12] Qin, Y., Zou, J., & Rhee, M. (2022). An AI-driven intelligent transaction execution framework for internet finance platforms: Cost-sensitive real-time integrity interdiction. Eurasia Journal of Science and Technology, 4(2), 36–44.

[13] Li, P., Ren, S., Zhang, Q., Wang, X., & Liu, Y. (2024). Think4SCND: Reinforcement learning with thinking model for dynamic supply chain network design. IEEE Access, 12, 195974–195985.

[14] Chen, J., Zhang, X., & Han, X. (2023). Explainable identification of tax-related financial reporting anomalies in U.S. public companies: Cross-fitted tax-consistency residuals and Owen-value Shapley propagation. World Journal of Economics and Business Research, 1(1), 33–45.

[15] Wang, J., Liu, J., Zheng, W., & Ge, Y. (2025). Temporal heterogeneous graph contrastive learning for fraud detection in credit card transactions. IEEE Access.

[16] Lin, Y., Wang, Y., & Bian, M. (2024). XOR-EW: An explainable and cost-aware artificial intelligence framework for operational risk early warning in financial institutions. AI and Data Science Journal, 1(1), 94–102.

[17] Liu, Y., Ren, S., Wang, X., & Zhou, M. (2024). Temporal logical attention network for log-based anomaly detection in distributed systems. Sensors, 24(24), 7949.

[18] Zou, J., Qin, Y., & Rhee, M. (2024). Beyond model scale: A task-aware reliability benchmark for large language models in clinical decision support. Innovation and Technology Studies, 1(1), 39–47.

[19] Lin, Y., Yue, X., Bian, M., Zhu, R., & Teng, D. (2026). Subgraph-level temporal graph transformer with laundering-motif prototypes for anti-money laundering on blockchains. IEEE Access.

[20] Xing, S., Wang, Y., & Liu, W. (2025). Multi-dimensional anomaly detection and fault localization in microservice architectures: A dual-channel deep learning approach with causal inference for intelligent sensing. Sensors, 25(11), 3396.

[21] Chen, J., Cui, Y., Zhang, X., Yang, J., & Zhou, M. (2024). Temporal convolutional network for carbon tax projection: A data-driven approach. Applied Sciences, 14(20), 9213.

[22] Fan, H., & Chen, J. (2025). From ledgers to large language models in contemporary accounting practice. American Journal of Big Data, 6(5).

[23] Liang, Y., & Jiao, Y. (2024). Real-time financial fraud detection using a feature-light temporal dynamic graph neural network with contrastive learning. Journal of Trends in Finance and Economics, 1(1), 76–84.

[24] Ping, W., & Fan, H. (2024). Disclosure quality and valuation risk in Level 3 investments: A disclosure-implied risk framework for U.S. business development companies. Social Science and Management, 1(1), 78–87.

[25] He, X., & Zhou, L. (2024). Adaheal: An adaptive framework for fault detection and automated recovery in large-scale cloud-native microservices. AI and Data Science Journal, 1(1), 132–141.

[26] Wang, Y., Bian, M., & Lin, Y. (2022). An explainable temporal graph neural network for disruption prediction and risk propagation in multi-tier supply chains. Journal of Computer Science and Electrical Engineering, 4(1), 17–25.

[27] Xing, S., & Wang, Y. (2022). MCG: Marginal–conditional graph learning for joint anomaly detection and root cause localization in distributed cloud systems. Journal of Computer Science and Electrical Engineering, 4(1), 81–88.

[28] Lu, Z., Zou, X., & Wei, K. (2022). Machine learning-based detection of BGP routing anomalies for enhanced network security and resilience: A hybrid multi-scale temporal learning approach. Journal of Computer Science and Electrical Engineering, 4(1), 62–71.

[29] Rhee, M., Zou, J., & Qin, Y. (2024). CLARITY: A closed loop framework for duplicate aware failure prediction and budgeted automated remediation in cloud infrastructure. AI and Data Science Journal, 1(1), 76–85.

[30] Chen, J., Liang, Y., Liu, J., & Zhou, M. (2026). Temporal transformer with conditional tabular GAN for credit card fraud detection: A sequential deep learning approach. Mathematics, 14(7), 1183.

[31] Mo, T., Zhang, C., Zou, J., Guo, Z., & Rhee, M. (2026). Self-evolving AI agents with dual memory for automated software testing and bug localization. IEEE Access.

[32] Wang, B., Wang, Z., Zhao, W., & Liu, Y. (2025). Network fabric simulation and validation for data center routing convergence under large-scale failure scenarios. Computer Science Bulletin, 8(1), 310–326.

[33] Yao, W. (2024). PrivTok-XR: Privacy-preserving real-time speech processing for accessible extended reality systems. Innovation and Technology Studies.

[34] Chen, J., & Fan, H. (2025). Beyond automation in tax compliance through artificial intelligence and professional judgment. Frontiers in Business and Finance, 2(2), 399–418.

[35] Fan, H., Jiao, Y., Wang, M., Chen, J., & Yang, S. (2026). Continual graph learning under strategic adversarial drift in dynamic networks. Scientific Reports.

[36] Zhou, L., & He, X. (2022). Dependency-aware root cause analysis and automated fault recovery for cloud-native microservices. Journal of Computer Science and Electrical Engineering, 4(1), 72–80.

[37] Wang, J., Ge, Y., Liu, J., & Jiang, B. (2024). Early warning of liquidity stress in U.S. banks with seasonally aligned frailty hazards and delayed-feedback conformal alarms. Journal of Trends in Finance and Economics, 1(1), 85–100.

[38] Qin, Y., Rhee, M., Teng, D., Zhang, C., & Zi, B. (2026). Neuro-symbolic constraint verification for LLM-driven internet finance transaction execution. Scientific Reports.

[39] Chen, J., Han, X., Zhang, X., & Yang, Y. (2023). Exact concept-level explanations for early detection of financial misstatements. Eurasia Journal of Science and Technology, 5(2), 34–48.

[40] Wang, Y., & Zhang, C. (2023). Aeris: Adaptive executor-scoped resource management for privacy-preserving data processing in Apache Spark. Eurasia Journal of Science and Technology, 5(2), 25–33.

[41] Chen, X. (2024). Who is accountable for AI decisions? Rethinking legal liability and ethical responsibility through counterfactual harm apportionment. Innovation and Technology Studies.

[42] Xu, Q., & Zhu, R. (2024). CoSTAR: Explainable and cost-sensitive machine learning for tax-aware corporate fraud risk detection. Innovation and Technology Studies.

[43] Zhang, F., & Wu, B. (2025). Large language models as general purpose intelligence systems for reasoning, planning and decision making. American Journal of Artificial Intelligence and Neural Networks, 6(4), 45–72.

[44] Xing, S., & Wang, Y. (2025). Proactive data placement in heterogeneous storage systems via predictive multi-objective reinforcement learning. IEEE Access, 13, 117986–117998.

[45] Zou, X., Wei, K., & Lu, Z. (2024). AEGIS: Calibrated anomaly-risk-driven checkpointing and proactive migration for fault-tolerant high-concurrency stream processing. Innovation and Technology Studies.

[46] Mo, T., Wang, G., Feng, E., & Li, P. (2023). Execution-feedback-guided context selection for large language model-based software testing and debugging. World Journal of Information and Knowledge Management, 1(1), 55–64.

[47] Xu, Q. (2023). Explainable machine learning for tax reporting anomaly detection and compliance risk assessment. World Journal of Economics and Business Research, 1(1), 56–66.

[48] Li, P., Bian, M., & Lin, Y. (2026). Carbon-aware offline reinforcement learning for joint inventory replenishment and transportation scheduling in multi-echelon supply chains. Highlights in Business, Economics and Management, 69, 251–262.

[49] Zhu, R., & Yue, X. (2024). An explainable machine learning framework for predicting dividend sustainability and financial risk of US REITs under interest-rate shocks. Journal of Trends in Finance and Economics, 1(2), 48–57.

[50] Li, P., Feng, E., Wang, G., & Mo, T. (2024). Hybrid deep reinforcement learning and integer optimization for resource scheduling in heterogeneous AI computing clusters. Innovation and Technology Studies, 1(1), 48–56.

[51] Ji, E., Wang, Y., Xing, S., & Jin, J. (2025). Hierarchical reinforcement learning for energy-efficient API traffic optimization in large-scale advertising systems. IEEE Access.

[52] Wang, Y., & Xing, S. (2025). AI-driven CPU resource management in cloud operating systems. Journal of Computer and Communications, 13(6), 135–149.

[53] Zhou, F. (2026). Spectral co-design of ship rudder geometry and heading autopilot for energy efficiency and manoeuvring safety. Transactions on Environment, Energy and Earth Sciences, 7, 158–181.

[54] Shen, A., & Huang, S. (2022). A systems-aware framework for intelligent large-scale recommendation models. Journal of Computer Science and Electrical Engineering, 4(1), 36–44.

[55] Zhang, F., & Yang, J. (2024). Robustness-aware model compression for multimodal perception under challenging conditions. AI and Data Science Journal, 1(1), 103–111.

[56] Chen, X. (2023). Algorithmic bias and discrimination in AI-driven lending: A Rashomon-based audit of less discriminatory alternatives and adverse-action accountability. World Journal of Sociology and Law.

[57] Yao, W. (2022). Raptor: Risk-adaptive security and privacy-preserving transaction processing for large-scale mobile payment systems. Journal of Computer Science and Electrical Engineering, 4(1), 89–97.

[58] Ren, S., Jin, J., Niu, G., & Liu, Y. (2025). ARCS: Adaptive reinforcement learning framework for automated cybersecurity incident response strategy optimization. Applied Sciences, 15(2), 951.

[59] Zhou, F., & Liu, W. (2026). Physics-anchored residual learning for ship fuel consumption and carbon dioxide emission prediction. Energy and AI, 100945.

[60] Liu, Y., Guo, L., Hu, X., & Zhou, M. (2025). Sensor-integrated inverse design of sustainable food packaging materials via generative adversarial networks. Sensors, 25(11), 3320.

[61] Zhang, F., & Yang, J. S. (2025). Learning driven decision intelligence for autonomous driving through multimodal understanding world modeling and policy optimization. Frontiers in Artificial Intelligence Research, 2(3), 616–634.

[62] Chen, X. (2026). Can people challenge an AI decision? Measuring the contestability chain in law and corporate disclosure. AI and Ethics, 6(6), 642.

[63] Sun, T., Yang, J., Li, J., Chen, J., Liu, M., Fan, L., & Wang, X. (2024). Enhancing auto insurance risk evaluation with transformer and SHAP. IEEE Access, 12, 116546–116557.