Proceedings of the 10th International Conference on Economic Management and Green Development
DOI: 10.54254/2754-1169/2026.37238

Quantitative Strategies and Liquidity Stress across Financial Crises: Evidence from U.S. Equities

Yanqi Jia

College Social Sciences, University of Birmingham, Birmingham, UK
YXJ596@student.bham.ac.uk

Abstract. To reconstruct the monthly S&P 500 membership from 2007 to 2025, this paper employs public records. This work is trying to minimize universe look-ahead bias of a sample limited to the constituents of a universe. The membership snapshot at the end of each month is used for the construction of the portfolio. It is not necessary that stocks have return histories for the whole sample. Licensed archives are more trustworthy than public records. To get a sense of the impact of this data constraint, this work compares point-in-time sample to a survivor panel using ex-post membership lists. Fixed weight assumptions for handling missing returns are also tested. This reconstruction affords an opportunity to focus on two components of the same universe: the performance of four equity strategies during normal months and crisis months, and the evolution of market illiquidity during the two periods. Regime differences in strategy performance. This work calculates the changes in log illiquidity by using a regression fit with HAC standard errors with six lags. Forecasts have moderate predictive power, but the rankings change, so it is not possible to rank the three models. Some sources of selection bias remain unresolved in the case of public reconstruction.

Keywords: quantitative strategies, financial crises, market liquidity, Amihud illiquidity, machine learning

1. Introduction

Volatility rises, funding constraints tighten, and order imbalances increase. During times such as these, strategy returns can turn very quickly, and the full-sample averages can hide the severity of the turn. Three strategies are based on past returns and the liquidity strategy is based on lagged ILLIQ. The cross-sectional momentum strategy is to buy stocks that have recently won and sell stocks that have recently lost [1]. Time-series momentum uses each asset's own return history to determine its position [2]. A strategy that exploits the negative serial correlation of individual stock returns is a short-term contrarian strategy [3]. They might vary in different crisis situations. Take into account the rebound from a panic selloff. Momentum strategies tend to lose money when market turns around [4]. Impaired liquidity, on the other hand, could change the return premium that investors receive for the stocks that aren't traded as easily.

But there's a liquidity dimension to these crisis outcomes. Trading costs impact the returns that investors demand [5] and the Amihud ratio capture the price change/dollar traded [6]. The literature

DOI: 10.54254/2754-1169/2026.37238

on asset-pricing assumes that the level of overall liquidity can serve as a priced state variable [7] and that required returns are related to expected liquidity and liquidity covariances [8]. Funding shocks can tighten margins and reduce market liquidity [9]. This work therefore examines changes in trading conditions separately from returns generated by the portfolio rules.

The empirical design incorporates three related tests. It compares the performance of time-series momentum, cross-sectional momentum, short-term reversal, and a liquidity-premium portfolio across normal months and two pre-specified crisis windows. It then estimates changes in market illiquidity and the liquidity premium during these windows. The final test examines whether month-end information predicts liquidity stress in the following month. Forecasting models include logistic regression, random forest, the bagged-tree method introduced by Breiman [10], and XGBoost, the regularized boosting system introduced by Chen and Guestrin [11]. Expanding-window evaluation preserves time order and addresses class imbalance without a random train-test split.

The results vary by regime. During the global financial crisis (GFC), annualized time-series and cross-sectional momentum returns are -18.8% and -27.7%. In the four-month COVID-19 window, they are 14.5% and 18.4%. The liquidity premium changes from 21.0% annualized in the GFC to -23.1% during COVID-19, while log market illiquidity rises in both crises. Under the baseline top-10% stress definition, random forest records the highest out-of-sample receiver operating characteristic area under the curve (ROC-AUC), 0.698. Logistic regression ranks first at the top-15% threshold, and random forest does so at the top-20% threshold. Thus, the model ordering depends on how stress is defined.

2. Related work

The four portfolio rules capture different forms of return predictability. Jegadeesh and Titman rank stocks by intermediate-horizon past performance and find a cross-sectional winner-minus-loser spread [1]. Moskowitz, Ooi, and Pedersen document time-series momentum in equity-index, currency, commodity, and bond futures [2]; an asset's own past return predicts its direction for up to about one year. These findings motivate the 12-to-2-month formation signal used here, but the two portfolios remain distinct. Cross-sectional momentum ranks stocks against one another. Time-series momentum determines each stock's trading direction from its own return history.

Short-term reversal is on a different time frame. However, Jegadeesh reports the occurrence of negative first order serial correlation (FOSC) for monthly stock returns [3]. This strategy is a short-term reversal strategy, which means purchasing stocks that did poorly last month, while selling stocks that did well last month. But this pattern with no conditions is not a complete explanation of the strategy during crises. Daniel and Moskowitz discover that the momentum strategies can experience large losses following declines [4], especially when the volatility is high and prices bounce back. Economically important state dependence may thus be identified in even short crisis windows.

The liquidity literature distinguishes trading frictions from compensation for liquidity risk. Amihud and Mendelson establish a relationship between expected returns and bid-ask spreads [5], and Amihud defines ILLIQ as the average daily absolute return divided by dollar trading volume [6]. Aggregate liquidity is examined in the context of exposure to innovations by Pastor and Stambaugh [7]. Acharya and Pedersen integrate expected liquidity and liquidity covariances into a single asset-pricing model [8], while Brunnermeier and Pedersen shed light on how market and funding liquidity can go hand in hand [9]. These studies have led us to investigate the effects of market illiquidity and strategy performance in times of crisis and to see if relative liquidity stress can be forecasted.

DOI: 10.54254/2754-1169/2026.37238

This study estimates ILLIQ, constructs an illiquid-minus-liquid portfolio, and uses relative stress as the prediction target. It compares three prediction models. Both tree-based models can capture interactions among the predictors. Because the monthly sample is small and stress events are rare, logistic regression serves as a benchmark for out-of-sample classification performance.

3. Data and methodology

The sample runs from 3 January 2007 through 31 December 2025. The S&P 500 represents the large-cap segment of the U.S. equity market [12]. Historical membership comes from a pinned public repository of dated component snapshots [13], whose README states that public change records were used to assemble the history. The resulting series is reproducible, but it is not an official licensed constituent archive. For every trading or formation date, the code carries forward the most recent snapshot dated on or before that date. Across 2007-2025, the sample includes 916 source tickers. Monthly membership ranges from 497 to 506 and has a median of 503. The design imposes no fixed stock count, sector draw, or full-period coverage screen. Yahoo Finance provides adjusted prices, raw closing prices, raw volume, and SPY [14]; Cboe provides VIX [15]. The frozen panel covers 4,780 trading days. An audit finds 245 source tickers with no usable Yahoo observations and six unresolved identities. They remain in the historical membership record rather than being dropped without disclosure.

Monthly stock returns are calculated from adjacent month-end adjusted closing prices. Daily adjusted returns are used for ILLIQ. Dollar volume equals the raw closing price multiplied by raw shares traded, which avoids combining a split-adjusted price with unadjusted volume. For stock i in month t, ILLIQ is the mean across valid days of |r_i,d| divided by dollar volume_i,d. At least 15 observations are required. Market ILLIQ is the cross-sectional median across stocks that are members in that month. Figure 1 scales this series by 10^10 for readability.

Signals use only information available at the end of month t and earn returns in t+1. For time-series momentum, the signal is the sign of each stock's cumulative return from t-12 through t-2; positive signals enter the long leg and negative signals enter the short leg. Cross-sectional momentum uses the same formation return but buys the top quintile and sells the bottom quintile. Short-term reversal buys the bottom quintile of month-t returns and sells the top quintile. The liquidity strategy buys the highest-ILLIQ quintile and sells the lowest-ILLIQ quintile. Rankings include only month-t members with the required signal. Each equally weighted leg is scaled to 0.5, giving unit gross exposure and intended net exposure of zero, and weights stay fixed through t+1. When a selected stock has no realized return, the baseline assigns 0%; sensitivity cases assign -30% and -100%. These are stated assumptions, not observed delisting returns, and missing positions do not cause the remaining weights to increase. The estimates exclude transaction costs and short-sale fees because consistent historical quote data are unavailable for the full sample.

The GFC window covers September 2008 through June 2009, and the COVID-19 window covers February through May 2020. All other valid months form the normal regime. Reported performance measures are annualized mean return, annualized volatility, their ratio as the Sharpe statistic for a self-financing portfolio, and maximum drawdown. Annualization multiplies the monthly mean by 12 and volatility by the square root of 12. Because the COVID-19 regime contains four observations, its annualized figures describe that window rather than a precise long-run premium.

To estimate crisis shifts, monthly log market ILLIQ is regressed on an intercept and indicators for the two fixed crisis windows. Heteroskedasticity-and-autocorrelation-consistent standard errors use six lags. Each coefficient is a log-point difference from normal months; exp(coefficient)-1 gives the

DOI: 10.54254/2754-1169/2026.37238

implied proportional change. The analysis reports the liquidity-premium return separately because market illiquidity and returns from holding illiquid stocks are different outcomes.

A month m is labeled as stressed when market ILLIQ exceeds the 90th percentile of the preceding 24 months. The threshold uses only observations before m and defines stress as the top 10% of the trailing distribution. Robustness checks lower the percentile to 85 and 80, producing top-15% and top-20% definitions while leaving all other inputs and estimation rules unchanged. Predictors dated m-1 include log market ILLIQ, mean VIX, annualized realized SPY volatility, the log change in aggregate dollar volume, the SPY return, and 12-to-2-month SPY momentum.

Logistic regression standardizes predictors and uses balanced class weights. Random forest has 150 trees, maximum depth four, minimum leaf size three, and balanced subsample weights. XGBoost has 150 depth-two trees, a 0.05 learning rate, 0.8 row and column subsampling, and a positive-class weight estimated inside each training sample. Every expanding window re-estimate scaling and class weights. Seven years of complete observations initialize the models, followed by one-month-ahead forecasts from January 2016 through December 2025. Each training month precedes the month it forecasts. Evaluation uses precision, recall, F1, and ROC-AUC. Accuracy is not reported because only 12 of the 120 baseline forecast months are stressed. Feature importance, estimated on the complete sample, serves only as a descriptive diagnostic and not as out-of-sample explanatory evidence.

4. Results

Table 1 reports regime-level performance for the four strategies in the primary point-in-time universe. In normal months, time-series momentum earns 0.5% annualized with a Sharpe statistic of 0.11. Cross-sectional momentum earns -0.3% with a Sharpe of -0.05, and short-term reversal earns 1.3% with a Sharpe of 0.22. The liquidity premium earns -0.5% annualized, with 4.7% volatility and a Sharpe statistic of -0.11. These estimates exclude implementation costs.

Table 1. Performance of four quantitative strategies by market regime

Strategy	Regime	N	Ann. ret.	Ann. vol.	Sharpe	Max DD
Time-series mom.	Normal	200	0.5%	4.6%	0.11	-14.5%
Time-series mom.	GFC	10	-18.8%	11.5%	-1.64	-16.7%
Time-series mom.	COVID-19	4	14.5%	14.4%	1.01	-3.7%
Cross-sectional mom.	Normal	200	-0.3%	7.4%	-0.05	-29.7%
Cross-sectional mom.	GFC	10	-27.7%	24.9%	-1.11	-30.4%
Cross-sectional mom.	COVID-19	4	18.4%	15.5%	1.19	-4.3%
Short-term reversal	Normal	212	1.3%	6.1%	0.22	-16.7%
Short-term reversal	GFC	10	-14.9%	12.3%	-1.21	-11.9%
Short-term reversal	COVID-19	4	-11.0%	24.0%	-0.46	-9.3%
Liquidity premium	Normal	213	-0.5%	4.7%	-0.11	-19.6%
Liquidity premium	GFC	10	21.0%	19.6%	1.07	-6.2%
Liquidity premium	COVID-19	4	-23.1%	16.6%	-1.39	-8.1%

DOI: 10.54254/2754-1169/2026.37238

During the GFC, annualized returns are -18.8% for time-series momentum, -27.7% for cross-sectional momentum, and -14.9% for short-term reversal. Cross-sectional momentum records a maximum drawdown of -30.4%. The illiquid-minus-liquid portfolio instead earns 21.0% annualized with 19.6% volatility. In the labeled ex-post survivor panel, the corresponding GFC return is 15.2%. That panel is a sensitivity check, not an unbiased alternative estimate. Nor does the positive spread imply easy execution: it relies on closing prices and omits crisis trading costs.

The ranking changes in the COVID-19 window. Time-series and cross-sectional momentum earn 14.5% and 18.4% annualized, whereas short-term reversal and the liquidity premium earn -11.0% and -23.1%. Four observations cannot establish stable COVID-era premia, but they do not fit a single crisis pattern. Across the 0%, -30%, and -100% missing-return assumptions, each scenario imputes six positions. The matching counts show that weights and missing identities remain fixed. The reproducibility files report the full set of scenario returns.

Figure 1 shows the largest market ILLIQ spike during the GFC and another sharp increase in March 2020, set against a long-run decline in the raw series. Table 2, panel A reports the regression estimates. The GFC coefficient is 1.236 (HAC standard error 0.121, $p < 0.001$), and the COVID-19 coefficient is 0.629 (standard error 0.103, $p < 0.001$). These log-point estimates imply increases of about 244% and 88% relative to normal months.

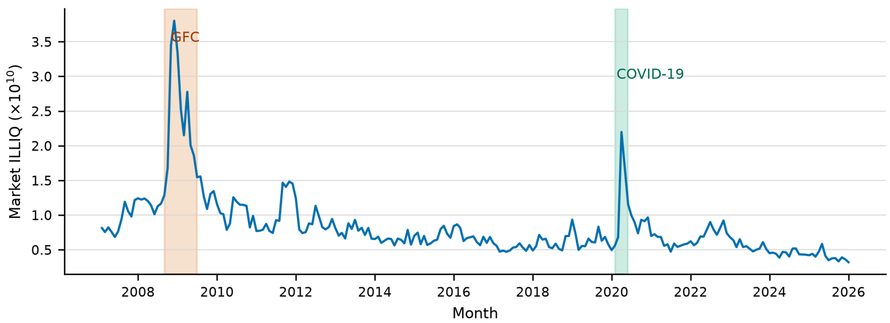

Figure 1. Market illiquidity from 2007 to 2025 with pre-specified crisis windows

Table 2. Crisis shifts in market illiquidity and the liquidity premium

Panel	Item	Estimate	SE	p	Ann. ret.	Ann. vol.	N
A. Market ILLIQ	GFC	1.236	0.121	<0.001			228
A. Market ILLIQ	COVID-19	0.629	0.103	<0.001			228
B. Liquidity premium	Normal				-0.5%	4.7%	213
B. Liquidity premium	GFC				21.0%	19.6%	10
B. Liquidity premium	COVID-19				-23.1%	16.6%	4

Table 2, panel B separates market illiquidity from the liquidity premium. The premium is -0.5% annualized in normal months, 21.0% in the GFC, and -23.1% during COVID-19. Illiquidity rises in both crises, but returns on the less-liquid quintile relative to the more-liquid quintile do not follow one pattern.

Table 3 reports 120 baseline forecasts, including 12 stress months. Logistic regression has precision 0.186, recall 0.667, F1 0.291, and ROC-AUC 0.678. Random forest records 0.375, 0.250,

DOI: 10.54254/2754-1169/2026.37238

0.300, and 0.698 on the same metrics. XGBoost records 0.333 for precision, recall, and F1, with ROC-AUC 0.620. Figure 2 shows the full out-of-sample ROC curves.

Table 3. Out-of-sample liquidity-stress classification performance

Model	N	Stress	Precision	Recall	F1	ROC-AUC
Logistic regression	120	12	0.186	0.667	0.291	0.678
Random forest	120	12	0.375	0.250	0.300	0.698
XGBoost	120	12	0.333	0.333	0.333	0.620

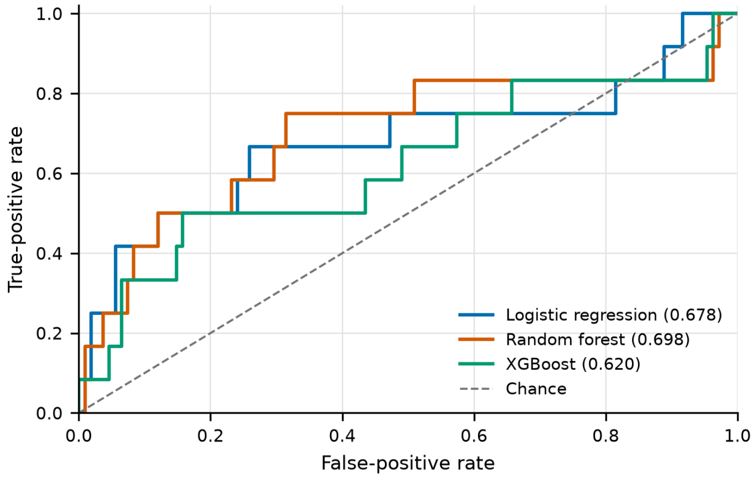

Figure 2. Out-of-sample receiver operating characteristic curves

Table 4 varies the stress threshold. The top-15% definition produces 14 stress months and ROC-AUC values of 0.737 for logistic regression, 0.706 for random forest, and 0.623 for XGBoost. The top-20% definition produces 19 stress months, with corresponding values of 0.746, 0.751, and 0.688. Random forest ranks first under the baseline and top-20% definitions; logistic regression ranks first under top-15% stress. Model performance therefore depends on the threshold, and the results do not support a general claim that nonlinear methods always add value.

Table 4. Sensitivity of ROC-AUC to the liquidity-stress definition

Definition	Stress	Logistic	Random forest	XGBoost
Top 10%	12	0.678	0.698	0.620
Top 15%	14	0.737	0.706	0.623
Top 20%	19	0.746	0.751	0.688

5. Discussion

Different results were obtained in the two crises using the same lagged trading rule. These estimates refer to the risk exposure due to fixed portfolio rules; they are not identifying a single causal mechanism.

DOI: 10.54254/2754-1169/2026.37238

The Market ILLIQ and the liquidity premium are two distinct concepts. ILLIQ measures the magnitude of price shocks per dollar of trading volume. The portfolio spread measures whether less liquid stocks outperform more liquid stocks later in the market. Changes in these two indicators are not necessarily synchronous; therefore, changes in market liquidity must be measured directly and cannot be inferred from a single indicator.

Due to the limited sample size, the forecasting results should be interpreted cautiously because only 12 of the 120 evaluation months are classified as stressed in their predictions. The best ROC-AUC exceeds the logistic-regression and second-best value by 0.02, and the model performance varied with changes in the stress threshold; no single model consistently outperformed in all situations.

No obvious use of future information in portfolio construction or prediction was found in the predicted sample. The constituent stock lists used to construct the portfolios are all dated before the month of construction; momentum metrics use observations up to month t, reversal and liquidity rankings use data from month t, and the returns for each portfolio are obtained in month t+1, with weights fixed before realization. Stress thresholds are determined solely by historical months, all predictors are dated m-1, and the cutoff month for each training set is earlier than the corresponding prediction month.

Data limitations are mainly reflected in three aspects. First, not all changes of constituents are recorded in publicly available data, and in some cases, the dates or the stock codes may not be correctly identified, and Yahoo Finance does not offer delisting returns like CRSP does. Second, the stock selection is done at the point in time and the absence of a full-sample screen reduces universe look-ahead bias, but is not able to remove the survivorship bias or selection bias, which is only addressed in the sensitivity analysis through archived survivor samples. Lastly, it is important to note that equal-weighted portfolios and closing prices ignore bid-ask spreads, market impact, borrowing constraints and turnover rates, and ILLIQ is merely an approximation of trading conditions.

6. Conclusion

The data in this study is based on the United States stock market data for the years 2007-2025. It discusses four quantitative approaches. The analysis includes monthly returns, levels of illiquidity in the market, and the predictability of liquidity pressures in the next month. There were differences in strategy performance between the two crisis windows. During the financial crisis, the market illiquidity index (ILLIQ) increased by some 1.24 log points. It increased by about 0.63 log points in the time of COVID-19. For prediction, the ROC-AUC (at the baseline of 10%) varied between 0.62 and 0.70. Random forest had the best performance at this threshold. The logistic regression models outperformed the other models when the threshold was raised to 15%. Random forest was again the best at the 20% threshold. Order of model performance varied between the different definitions of stress.

The results do not establish one strategy as dominant across crises, nor do they show that machine learning improves every forecast. Point-in-time reconstruction is based on public data and, as such, is less comprehensive than a licensed CRSP style database in terms of constituent history, mapping ticker changes, and handling missing returns. Future research may investigate these restrictions by incorporating confirmed delisting returns, measures of the liquidity of trading in each transaction, explicit measures of trading costs and, of course, evidence from more crisis episodes.

DOI: 10.54254/2754-1169/2026.37238

References

[1] Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance, 48*(1), 65-91. https://doi.org/10.1111/j.1540-6261.1993.tb04702.x
[2] Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics, 104*(2), 228-250. https://doi.org/10.1016/j.jfineco.2011.11.003
[3] Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *The Journal of Finance, 45*(3), 881-898. https://doi.org/10.1111/j.1540-6261.1990.tb05110.x
[4] Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics, 122*(2), 221-247. https://doi.org/10.1016/j.jfineco.2015.12.002
[5] Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics, 17*(2), 223-249. https://doi.org/10.1016/0304-405X(86)90065-6
[6] Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets, 5*(1), 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6
[7] Pastor, L., & Stambaugh, R. F. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy, 111*(3), 642–685. https://doi.org/10.1086/374184
[8] Acharya, V. V., & Pedersen, L. H. (2005). Asset pricing with liquidity risk. *Journal of Financial Economics, 77*(2), 375–410. https://doi.org/10.1016/j.jfineco.2004.06.007
[9] Brunnermeier, M. K., & Pedersen, L. H. (2009). Market liquidity and funding liquidity. *The Review of Financial Studies, 22*(6), 2201–2238. https://doi.org/10.1093/rfs/hhn098
[10] Breiman, L. (2001). Random forests. *Machine Learning, 45*, 5–32. https://doi.org/10.1023/A:1010933404324
[11] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). https://doi.org/10.1145/2939672.2939785
[12] S&P Dow Jones Indices. (n.d.). S&P U.S. indices methodology. https://www.spglobal.com/spdji/en/methodology/article/sp-us-indices-methodology/
[13] fja05680. (n.d.). S&P 500 historical components and changes (commit c31ac3cc56f28cf9a02b4e694eff7ceab596a0ff) [Data set]. MIT License. Retrieved August 22, 2026, from https://github.com/fja05680/sp500/tree/c31ac3cc56f28cf9a02b4e694eff7ceab596a0ff
[14] Yahoo Finance. (n.d.). Historical market data. Retrieved August 22, 2026, from https://finance.yahoo.com/
[15] Cboe Global Markets. (n.d.). VIX index historical data. Retrieved August 22, 2026, from https://www.cboe.com/tradable_products/vix/vix_historical_data