Predicting Stock Market Liquidity: A Multiple Regression Analysis

Cornelius Benle

Anderson Natasha

December 16, 2025

Mini-Project

1. Introduction

Stock market liquidity refers to how quickly a stock can be bought or sold without causing large changes in its price. Liquidity is essential for efficient markets because it minimizes transaction costs, enhances market stability, and supports price discovery.

This project examines whether trading activity and firm characteristics can explain variation in liquidity. Understanding these relationships assists investors, market makers, and regulators in evaluating liquidity conditions.

Research Question: Can stock liquidity (measured by average time between trades) be predicted using market activity and firm-level financial characteristics?

Objectives:

1. To build and evaluate a multiple linear regression model for stock liquidity.
2. To quantify the associations between liquidity and key predictors (VOLUME, NTRAN, PRICE, VALUE, DEBEQ).
3. To test the statistical significance of each predictor.
4. To assess model fit using residual diagnostics and RMSE.
5. To interpret findings using financial theory and liquidity literature.

2. Data and Methods

Dataset

The dataset contains 124 NYSE/AMEX-listed firms obtained from the University of Wisconsin-Madison ("Liquidity.csv"). The variables include:

Table 1: *

Variable Descriptions

Variable	Description
AVGT	Average time between trades; the liquidity measure used as the response variable(In minutes).
NTRAN	Number of transactions for the stock during the period.
VOLUME	Total number of shares traded.
PRICE	Stock price during the observation period.
VALUE	Market value (firm size), calculated as price times shares outstanding.
DEBEQ	Debt-to-equity ratio capturing firm leverage.

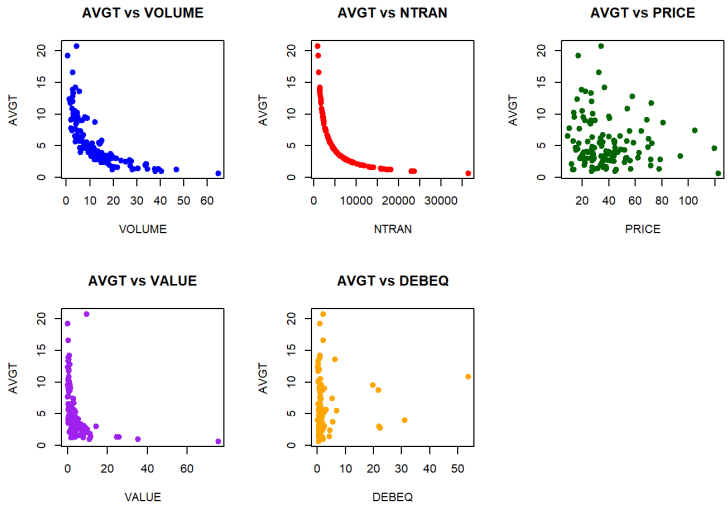

Figure 1: Scatter plot of AVGT and predictors

Analytical Procedure

Exploratory analysis showed strong skewness and an exponential relationship between AVGT and the predictors. Hence an exponential regression model was used as the base model and a log-linear regression model was used as the final regression model .

base model: $y = \beta_0 * e^{\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon}$

final model: $log(y) = log(\beta_0) + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon.$

Predictive accuracy was assessed using RMSE . $RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$

3. Results

Regression Output (Final Model)

Table 2: *
Final Regression Model: $\log(\text{AVGT}) \sim \text{NTRAN} + \text{VALUE} + \text{VOLUME} + \text{PRICE}$

Coefficient	**Estimate**	**Std. Error**	**t value**	**Pr(¿—t—)**
Intercept	2.507e+00	6.426e-02	39.008	$< 2 \times 10^{-16}$
NTRAN	-1.207e-04	1.165e-05	-10.360	$< 2 \times 10^{-16}$
VALUE	3.731e-02	4.925e-03	7.576	8.79×10^{-12}
VOLUME	-2.086e-02	5.258e-03	-3.967	1.25×10^{-4}
PRICE	-3.648e-03	1.244e-03	-2.933	0.004035
Residual Std. Error	0.2512 on 118 df			
Multiple R-squared	0.8757			
Adjusted R-squared	0.8715			
F-statistic	207.8 on 4 and 118 df		p-value: $< 2.2 \times 10^{-16}$	

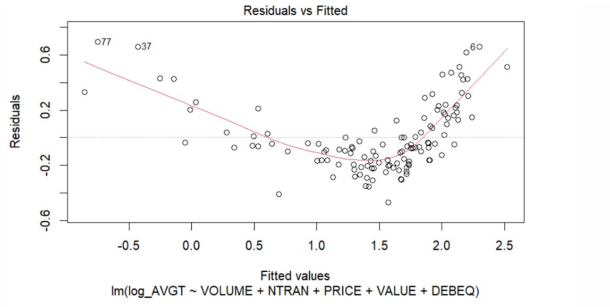

Figure: Residual Plot

Figure 2: *
Figure: Residual Plot

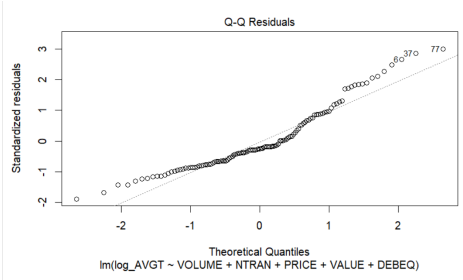

Figure 3: *
Figure: Q–Q Plot of Residuals

Model Accuracy (RMSE)

$$\boxed{RMSE = 1.946575}$$

The RMSE of 1.9466 indicates that, on average, our predictions deviate from the true AVGT(Average time between trades) values by nearly 2 minutes.

4. Discussion

The final log-linear regression model indicates that **NTRAN, VALUE, VOLUME, and PRICE are statistically significant predictors of stock liquidity**. These findings are consistent with results reported in the existing literature, which also identifies trading activity and firm size as major determinants of liquidity performance. In contrast, **DEBEQ was not statistically significant**, aligning with prior research suggesting that leverage ratios do not play a major role in short-term trading dynamics.

The estimated coefficients reveal that **NTRAN, VOLUME, and PRICE exhibit negative relationships with liquidity**, meaning that increases in these variables are associated with lower average time between trades (i.e., faster trading). This is theoretically intuitive: higher transaction counts and trading volume typically enhance market activity, while stock price levels may reflect differing trading intensities.

Because the model is log-linear, the interpretation of each coefficient is multiplicative rather than additive. Specifically, a one-unit increase in predictor x_i results in a multiplicative change of e^{β_i} in AVGT. Thus, each significant predictor influences liquidity by proportionally scaling the expected average time between trades.

Residual diagnostics also reveal important model limitations. The residual plot shows signs of **heteroscedasticity**, and the Q–Q plot indicates noticeable departures from normality. .

5. Conclusion

Although the model explains 87.15% of the variation in liquidity, the RMSE of 1.946 indicates a relatively high level of prediction error. Furthermore, the residual diagnostics reveal violations of important regression assumptions, including heteroscedasticity and non-normality. These issues raise concerns about the model's appropriateness for accurate prediction. For this reason, we recommend exploring nonlinear models or machine learning approaches and comparing their predictive accuracy with that of the regression model.

References

[Alp .Alp .2022] alp2022liquidityAlp, S. . 2022. Stock Prices and Their Impact on Stock Market Liquidity Stock prices and their impact on stock market liquidity. International Journal of Finance and Economics.

[AmihudAmihud2002] amihud2002illiquidityAmihud, Y. 2002. Illiquidity and stock returns: cross-section and time-series effects Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets5131–56.

[AriesantiAriesanti2015] ariesanti2015liquidityAriesanti, D. 2015. The Effect of Stock Prices and Trading Volume on Stock Liquidity The effect of stock prices and trading volume on stock liquidity. Indonesian Capital Market Review.

[FreesFrees2010] freesdataFrees, EW. 2010. Regression modeling with actuarial and financial applications: Data website. Regression modeling with actuarial and financial applications: Data website. https://instruction.bus.wisc.edu/jfrees/jfreesbooks/Regression Accessed February 14, 2025

[LesmondLesmond2005] lesmond2005liquidityLesmond, DA. 2005. Liquidity of emerging markets Liquidity of emerging markets. Journal of Financial Economics772411–452.