

The Role of Algorithmic Trading in the U.S. Stock Market

Z. Zhang

ABSTRACT

This paper reviews empirical research on algorithmic trading in the U.S. stock market and argues that its effects on price efficiency, transaction costs, and market volatility all stem from a single underlying behavior: the voluntary supply of liquidity by algorithmic and high-frequency trading firms. These firms are not required to keep trading when conditions turn unfavorable, so they supply liquidity when doing so is profitable and withdraw it when it is not. Under ordinary market conditions, this behavior narrows bid-ask spreads, speeds up the incorporation of new information into prices, and lowers trading costs for participants who can move as quickly as the market does. The same behavior also explains why these benefits break down during rare, correlated market shocks: firms that reliably supply liquidity on calm days can withdraw from the market in unison once conditions become risky, an event best illustrated by the Flash Crash of May 6, 2010. Bringing these three strands of research together shows that algorithmic trading has not simply made U.S. equity markets more efficient or more fragile. Instead, it has tied everyday market quality to a form of liquidity supply that is dependable under normal conditions and prone to sudden, collective withdrawal exactly when markets need it most.

KEYWORDS

algorithmic trading, high-frequency trading, liquidity provision, market efficiency, transaction costs, market volatility

HOW TO CITE THIS ARTICLE

Z. Zhang (2026). The Role of Algorithmic Trading in the U.S. Stock Market. *The Journal of Young Innovators*, 2(2), pp. 16–21.

DOI <https://doi.org/10.67419/jyi.v2i2.14>

READ ONLINE <https://young-innovator.org/issues/articles/algorithmic-trading>

Introduction

Most orders in the U.S. stock market are now generated and routed by computer programs that react to new information within milliseconds (Menkveld, 2016; O'Hara, 2015). A large share of this activity comes from high-frequency trading firms, which earn their profits by quoting prices at which they are willing to buy and sell and updating those quotes. Most high-frequency liquidity providers differ from the designated market makers of earlier decades in one important way: they are not formally obligated to keep trading when conditions turn unfavorable. Finance research calls this arrangement endogenous liquidity provision, since these firms supply continuous quotes because doing so is profitable, not because a rule requires it (Anand & Venkataraman, 2016). This distinction connects three literatures that are usually reviewed on their own: work on price efficiency, on transaction costs, and on market volatility.

This review argues that one behavior, the voluntary and conditional supply of liquidity by algorithmic and high-frequency traders, accounts for the findings in all three areas. Under ordinary conditions, this liquidity is supplied continuously, which improves price efficiency and lowers trading costs. Under the rare, correlated shocks that occasionally strike U.S. equity markets, the same liquidity can vanish just when it is needed most, producing the kind of breakdown seen during the Flash Crash of May 6, 2010 (Kirilenko et al., 2017).

This review synthesizes peer-reviewed empirical studies of algorithmic and high-frequency trading in U.S. equity markets, along with related studies of futures and foreign exchange markets used for comparison, published between 1970 and 2022.

Literature Review

Price Efficiency

Continuous liquidity provision is what allows algorithmic trading to improve price efficiency in Fama's (1970) sense of how quickly prices incorporate available information. A liquidity provider that constantly updates its quotes is, by construction, constantly folding new information into the price it is willing to trade at. Several studies confirm that this activity narrows the gap between market prices and their information-based value. Hendershott et al. (2011) use the New York Stock Exchange's 2003 move to automated quote dissemination as a natural experiment and find that algorithmic trading narrowed bid-ask spreads and made quoted prices more informative for large stocks, the same stocks where algorithmic

quoting is most active. Brogaard et al. (2014) reach a similar conclusion using a different method: high-frequency traders' own trades push prices toward their permanent, information-based level, a pattern that holds even on a stock's most volatile days. Carrion (2013) shows the mechanism at work, finding that high-frequency traders on NASDAQ supply liquidity when it is scarce and withdraw it when it is abundant, and that prices incorporate new information most efficiently on the days when their participation is highest. Chaboud et al. (2014) find a similar pattern in foreign exchange trading, where algorithmic liquidity supply is associated with fewer arbitrage opportunities and less predictable short-term returns.

This liquidity does not come for free. Chaboud et al. (2014) find that algorithms also take liquidity, not only supply it, and that much of the reduction in arbitrage opportunities comes from algorithms trading against slower counterparties. The firms that make prices more efficient are the same firms being paid for it, which is the subject of the next section.

Transaction Costs

If algorithmic traders profit from supplying liquidity, then part of the cost of trading is simply the price other participants pay for that liquidity. Beyond any commission, a trader pays a cost built into the bid-ask spread and the price impact of a trade, and Engle et al. (2012) point to a third cost: time. An order that must be filled right away typically costs more per share than one spread out over the day, since it demands liquidity instead of waiting for it to appear. Engle et al. (2012) model this tradeoff between cost and risk and show that algorithms designed to minimize cost, such as those that spread an order in proportion to volume, achieve lower average costs than urgent orders, because they trade on the liquidity provider's terms.

This arrangement rewards speed; Hendershott et al. (2011) find that narrower spreads translate into lower execution costs, especially for large, liquid stocks, while traders who cannot react as quickly continue to pay more of the remaining cost. Boehmer et al. (2021) examine 42 equity markets and find that algorithmic trading lowers the execution shortfalls realized by institutional investors overall, consistent with cheaper access to liquidity where algorithmic supply is present. Brogaard et al. (2014), using a natural experiment based on latency-reducing technology upgrades on the London Stock Exchange, find no clear relationship between high-frequency trading activity and the costs institutional investors actually paid. This result fits the same pattern: investors who cannot match the speed of algorithmic liquidity providers do not automatically share in the average savings that faster

traders capture. Lower transaction costs, then, are less a universal discount than a return paid to whoever can supply or access liquidity fastest.

Market Volatility

The same liquidity providers who narrow spreads and speed up price discovery on ordinary days are not obligated to keep supplying liquidity once conditions turn unfavorable, and this conditional supply is what drives the empirical record on volatility. Under ordinary conditions, continuous liquidity provision dampens volatility. Chaboud et al. (2014) find no evidence that algorithmic trading raises exchange-rate volatility; if anything, busier algorithmic periods are slightly calmer, because more liquidity is available to absorb everyday price swings. Hasbrouck and Saar (2013) find the same pattern in U.S. equities: more low-latency activity is associated with lower short-term volatility, tighter spreads, and a deeper limit order book, even during a period of falling prices. Menkveld (2016) reaches a similar conclusion in a review of the broader literature, finding that on an ordinary trading day, algorithmic and high-frequency liquidity supply tends to help rather than hurt market quality.

This supply is not guaranteed, and a growing body of research documents what happens when it disappears. Anand and Venkataraman (2016) use audit-trail data from the Toronto Stock Exchange and find that market makers scale back their liquidity provision together when conditions turn unfavorable, exactly what an unobligated liquidity provider would do once trading grows riskier. Kirilenko et al. (2017) document this withdrawal during the Flash Crash of May 6, 2010: the high-frequency firms that had been supplying liquidity under calm conditions continued trading as prices collapsed, rather than stepping in to absorb the imbalance. Brogaard et al. (2018) find that this withdrawal depends on the size of the shock. When a single stock has an extreme price movement, high-frequency traders keep supplying liquidity and absorb the shock; when many stocks move sharply together, these same traders demand more liquidity than they supply, and their stabilizing role reverses at the moment a market-wide shock begins. Boehmer et al. (2021) measure volatility across 42 countries, and their measure spans both ordinary and stressed periods together. They find that algorithmic trading raises short-term volatility on average, a result that fits this pattern rather than contradicting it: liquidity that is reliably supplied on calm days and reliably withdrawn during shocks will raise average volatility once both kinds of days are counted together.

Discussion

Algorithmic and high-frequency trading firms act as endogenous liquidity providers, supplying continuous, competitively priced liquidity when doing so is profitable and withdrawing it, often in unison, once conditions make continued supply unprofitable or risky (Anand & Venkataraman, 2016; Brogaard et al., 2018). This single behavior accounts for the efficiency gains documented by Hendershott et al. (2011) and Brogaard et al. (2014), the cost savings documented by Engle et al. (2012) and Boehmer et al. (2021), and the pattern of calm-market stability and crisis-period fragility found across the volatility literature. It also explains why these benefits are not shared evenly: liquidity that is sold rather than given away is priced according to speed, which is why slower participants capture less of the savings that faster ones do (Brogaard et al., 2014).

The Flash Crash of May 6, 2010, is evidence that this arrangement carries a specific, foreseeable risk. Because algorithmic liquidity provision is voluntary rather than obligated, it is available when it is easy to supply and prone to vanish just when it would matter most (Kirilenko et al., 2017). Better technology alone will not fix this, since the withdrawal documented by Anand and Venkataraman (2016) and Brogaard et al. (2018) is a rational response to unfavorable conditions rather than a malfunction. Addressing it means changing the incentives liquidity providers face during periods of stress, or else accepting synchronized withdrawal as a predictable cost of relying on voluntary rather than obligated liquidity supply.

Some proposed remedies target the market design that rewards speed itself rather than the withdrawal behavior. Budish et al. (2015) argue that the continuous limit order book rewards raw speed over the value of a trade, and propose frequent batch auctions that would reduce the advantage of being the fastest liquidity provider rather than simply the most willing one. Aquilina et al. (2022) estimate that the split-second races this speed advantage creates account for roughly one-fifth of total trading volume in the market they studied, a cost ultimately borne by slower liquidity providers and, indirectly, by the investors who trade with them. Market-wide circuit breakers reflect a related idea: since liquidity withdrawal is a rational response to unfavorable conditions rather than a rule violation, oversight has increasingly focused on containing its consequences rather than preventing it outright (U.S. Securities and Exchange Commission, 2016).

This review has traced one behavior, the voluntary supply of liquidity by algorithmic and high-frequency traders, through the literature on price efficiency, transaction costs, and market volatility. Continuous liquidity supply makes markets more efficient and cheaper to

trade in day-to-day, and its voluntary character makes markets vulnerable to synchronized withdrawal during rare, correlated shocks. The clearest evidence for this pattern remains the Flash Crash of May 6, 2010, when the same firms that had stabilized the market in calm conditions did not change their behavior as it came apart (Kirilenko et al., 2017).

References

- Anand, A., & Venkataraman, K. (2016). Market conditions, fragility, and the economics of market making. *Journal of Financial Economics*, 121(2), 327–349. <https://doi.org/10.1016/j.jfineco.2016.03.006>
- Aquilina, M., Budish, E., & O'Neill, P. (2022). Quantifying the high-frequency trading “arms race.” *The Quarterly Journal of Economics*, 137(1), 493–564. <https://doi.org/10.1093/qje/qjab032>
- Boehmer, E., Fong, K., & Wu, J. (2020). Algorithmic trading and market quality: International evidence. *Journal of Financial and Quantitative Analysis*, 56(8), 2659–2688. <https://doi.org/10.1017/S0022109020000782>
- Brogaard, J., Carrion, A., Moyaert, T., Riordan, R., Shkilko, A., & Sokolov, K. (2018). High frequency trading and extreme price movements. *Journal of Financial Economics*, 128(2), 253–265. <https://doi.org/10.1016/j.jfineco.2018.02.002>
- Brogaard, J., Hendershott, T., Hunt, S., & Ysusi, C. (2014). High-frequency trading and the execution costs of institutional investors. *The Financial Review*, 49(2), 345–369.
- Brogaard, J., Hendershott, T., & Riordan, R. (2014). High-frequency trading and price discovery. *The Review of Financial Studies*, 27(8), 2267–2306. <https://doi.org/10.1093/rfs/hhu032>
- Budish, E., Cramton, P., & Shim, J. (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *The Quarterly Journal of Economics*, 130(4), 1547–1621. <https://doi.org/10.1093/qje/qjv027>
- Carrion, A. (2013). Very fast money: High-frequency trading on the NASDAQ. *Journal of Financial Markets*, 16(4), 680–711. <https://doi.org/10.1016/j.finmar.2013.06.005>
- Chaboud, A. P., Chiquoine, B., Hjalmarsson, E., & Vega, C. (2014). Rise of the machines: Algorithmic trading in the foreign exchange market. *The Journal of Finance*, 69(5), 2045–2084. <https://doi.org/10.1111/jofi.12186>
- Engle, R., Ferstenberg, R., & Russell, J. R. (2012). Measuring and modeling execution cost and risk. *The Journal of Portfolio Management*, 38(2), 14–28. <https://doi.org/10.3905/jpm.2012.38.2.014>

- Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), 383–417. <https://doi.org/10.1111/j.1540-6261.1970.tb00518.x>
- Hasbrouck, J., & Saar, G. (2013). Low-latency trading. *Journal of Financial Markets*, 16(4), 646–679. <https://doi.org/10.1016/j.finmar.2013.05.003>
- Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *The Journal of Finance*, 66(1), 1–33. <https://doi.org/10.1111/j.1540-6261.2010.01624.x>
- Kirilenko, A., Kyle, A. S., Samadi, M., & Tuzun, T. (2017). The flash crash: High-frequency trading in an electronic market. *The Journal of Finance*, 72(3), 967–998. <https://doi.org/10.1111/jofi.12498>
- Menkveld, A. J. (2016). The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics*, 8, 1–24. <https://doi.org/10.1146/annurev-financial-121415-033010>
- O'Hara, M. (2015). High-frequency market microstructure. *Journal of Financial Economics*, 116(2), 257–270. <https://doi.org/10.1016/j.jfineco.2015.01.003>
- U.S. Securities and Exchange Commission. (2016, March 3). Testimony on regulatory reforms to improve equity market structure. <https://www.sec.gov/news/testimony/testimony-regulatory-reforms-to-improve-equity-market-structure.html>