

The Mathematics of Slippage: When Latency Arbitrage Stops Paying

Boris Fesenko*

April 2026

Abstract

Latency arbitrage in over-the-counter foreign exchange has been a profitable retail and small-institutional strategy for over two decades. Its profitability rests on three measurable quantities: the trader’s latency edge ε relative to a slower counterparty, the volatility of the price process between signal and execution, and the cost structure imposed by the broker. We formalize the relationship between these quantities and derive a closed-form break-even condition for the strategy. We then extend the model to N competing latency arbitrageurs and show that the equilibrium dispersion of latencies determines the long-run profit margin available to any individual participant. Calibrating to retail EUR/USD tick data for 2025–2026, the break-even latency floor τ^* sits in the 5–25 millisecond range depending on volatility regime, and the addition of a fourth competitor at equal latency reduces individual expected profit by approximately 75% relative to the monopoly case under our baseline calibration. The paper also discusses how broker-side execution adjustments — the “anti-arbitrage” plugin family — interact with the model and shift the effective break-even latency upward by roughly 50% in typical retail conditions. We conclude with practical implications and open research questions.

1. Introduction

Latency arbitrage exploits a transient mispricing between two correlated price feeds. In its simplest retail form, a trader maintains a fast reference feed (e.g., a tier-one ECN aggregator or a major-bank prime broker) and a slower broker feed, and sends orders against the broker feed at moments when the reference has moved but the broker has not yet repriced. The strategy was widely profitable on retail FX in the 2000s and early 2010s. By the mid-2020s, its profit margin has narrowed substantially. Practitioners and academic observers attribute the compression to two factors: (i) brokers have invested heavily in execution-side defenses, collectively documented in the practitioner literature (Fesenko, 2026a); and (ii) the latency competition between participants has compressed the size of the available edge.

This paper formalizes both effects in a single model and derives the conditions under which the strategy ceases to be profitable for an individual participant. Our contribution is threefold.

*BJF Trading Group Inc., Ontario, Canada. Email: boris@fesenko.ca. The author is the founder of BJB Trading Group and has developed retail and small-institutional latency-arbitrage software since 2000. The empirical calibration in Section 5 draws on data and instrumentation from those tools, but the model and analytical results are the author’s own.

First, we provide a closed-form decomposition of the expected per-trade profit into three components: a deterministic latency edge against a slower counterparty, a stochastic slippage cost during the trader’s own execution window, and the broker’s transaction costs.

Second, we characterize the break-even latency τ^* — below which the strategy generates positive expected profit — as a function of volatility, signal arrival rate, and per-trade fixed costs. The threshold is monotonically increasing in volatility and decreasing in signal arrival rate, with closed-form low- and high-intensity asymptotics.

Third, we extend the model to a Bertrand-style latency competition among N participants and derive the implied profit dispersion as a function of competitor count and latency dispersion.

We then calibrate the model to retail EUR/USD tick data for 2025–2026, illustrating where the break-even threshold sits empirically and how it has shifted relative to a 2010–2015 baseline. The empirical exercise is illustrative rather than definitive — a formal estimation on a multi-broker tick panel is left to future work — but suggests that the break-even latency on retail EUR/USD has compressed by approximately a factor of four over the past decade.

The remainder of the paper is organized as follows. Section 2 reviews related literature. Section 3 sets up the model. Section 4 derives the profitability conditions. Section 5 provides a numerical illustration. Section 6 discusses implications, including the interaction with broker-side defenses. Section 7 concludes.

2. Related Literature

Three strands of literature inform this work.

The first is the empirical microstructure literature on high-frequency trading and latency races. Hendershott, Jones and Menkveld (2011) document the price-discovery contributions of high-frequency trading in U.S. equities; Brogaard, Hendershott and Riordan (2014) extend the analysis to news-driven episodes. Menkveld (2013, 2016) provides foundational work on the economics of high-frequency trading and the dispersion of latency-driven returns. Aquilina, Budish and O’Neill (2022) quantify the welfare cost of the latency arms race and estimate the magnitude of liquidity transfers it generates — their estimate of approximately five basis points in the U.S. equity market is suggestive but not directly applicable to retail FX.

The second strand is the theoretical literature on latency competition and frequent-batch auctions. Budish, Cramton and Shim (2015) propose discrete-time auctions as a remedy to continuous-time latency races and provide a tractable model of the stale-quote sniping problem; their setup is the closest theoretical antecedent to the model below. Foucault, Hombert and Roşu (2016) model speed in news trading and derive the equilibrium return to speed under endogenous information acquisition.

The third strand is the FX market microstructure literature. Stoll (1989, 2000) lays groundwork on transaction costs; Hasbrouck (2018) provides empirical methodology for tick-level cost estimation;

Easley, López de Prado and O'Hara (2012) introduce the VPIN measure of order-flow toxicity, which we adapt indirectly in Section 5 by treating the per-second event rate λ as an empirical proxy for short-horizon information flow.

Our model differs from the cited frequent-batch auction literature in two ways: it is set in continuous time, and it is calibrated to retail FX rather than equity venues. The closest precedent in spirit is Foucault, Kadan and Kandel (2013) on liquidity cycles, although our focus is on the trader's profitability rather than market-maker incentives.

3. Model

3.1 Price process and information events

Let the fundamental price of a currency pair (we use EUR/USD as the leading example) follow a continuous-time process. Information events arrive according to a homogeneous Poisson process with intensity λ . Conditional on an event at time t_0 , the fundamental price jumps by ΔP , where ΔP is symmetric with $\mathbb{E}[|\Delta P|] = \mu_J$ and $\text{var}[|\Delta P|] = \sigma_J^2$.

Between events, the price evolves as a martingale with diffusion coefficient σ_D , so that for an event-free window of length τ :

$$\mathbb{E}[P_{t+\tau} - P_t \mid \text{no event}] = 0, \quad \text{var}[P_{t+\tau} - P_t \mid \text{no event}] = \sigma_D^2 \tau.$$

Over a horizon τ that includes the (random) Poisson events, the unconditional variance is:

$$\sigma^2(\tau) = \sigma_D^2 \tau + \lambda \sigma_J^2 \tau = (\sigma_D^2 + \lambda \sigma_J^2) \tau.$$

We denote the combined coefficient $\sigma^2 \equiv \sigma_D^2 + \lambda \sigma_J^2$ in what follows.

3.2 Latency competition

Let the trader have own latency τ_{self} , defined as the time between an information event reaching the public reference feed and the trader's order arriving at the broker's matching engine. Denote by τ_{broker} the latency of the broker's own pricing path (LP feed $\rightarrow$ bridge $\rightarrow$ quote engine), and assume τ_{broker} is fixed and known.

The strategy is profitable on a given event if and only if $\tau_{\text{self}} < \tau_{\text{broker}}$, i.e., the trader can act on the information before the broker reprices. The latency edge is:

$$\varepsilon \equiv \tau_{\text{broker}} - \tau_{\text{self}} > 0.$$

If τ_{broker} is fixed and the trader is the unique arbitrageur, the per-event gross profit is $q \cdot |\Delta P|$ with order size q , realized whenever $\varepsilon > 0$.

3.3 Slippage decomposition

Slippage is the difference between the price at which the trader intended to execute and the price at which the order actually fills. For a market order with own latency τ_{self} , slippage decomposes naturally into a continuous (diffusion) component and a discrete (jump) component:

$$S = S_{\text{drift}} + S_{\text{jump}}.$$

Diffusion component. The price drifts with diffusion coefficient σ_D during the order's transit. Modelling the drift as a Brownian motion stopped at τ_{self} , the absolute value $|S_{\text{drift}}|$ is a folded normal random variable. Its expectation equals:

$$\mathbb{E}[|S_{\text{drift}}|] = \sigma_D \sqrt{\frac{2\tau_{\text{self}}}{\pi}}.$$

Jump component. With probability $\lambda\tau_{\text{self}}$ a new event arrives during the order's transit, in which case the additional slippage equals $|\Delta P|$. Higher-order events (two or more jumps in the same transit window) carry probability $O((\lambda\tau_{\text{self}})^2)$ and we neglect them in the leading-order expression. Hence:

$$\mathbb{E}[S_{\text{jump}}] \approx \lambda\tau_{\text{self}}\mu_J.$$

Total slippage. Combining:

$$\mathbb{E}[S] \approx \sigma_D \sqrt{\frac{2\tau_{\text{self}}}{\pi}} + \lambda\tau_{\text{self}}\mu_J. \quad (\star)$$

For small τ_{self} the diffusion term dominates ($O(\sqrt{\tau})$). For moderate τ_{self} and high-event-rate windows the linear jump term dominates ($O(\tau)$); the crossover occurs at $\tau_{\text{self}} \approx 2\sigma_D^2/(\pi\lambda^2\mu_J^2)$.

4. Profitability Conditions

4.1 Single-arbitrageur break-even

Define per-trade gross profit conditional on winning the latency race (i.e., $\tau_{\text{self}} < \tau_{\text{broker}}$) as:

$$\pi_{\text{gross}} = q \cdot |\Delta P| - q \cdot S - C,$$

where C denotes the fixed per-trade cost (commission plus the spread cost on the entry leg, in monetary units).

Conditional on an event and on winning the race, expected per-trade profit is:

$$\mathbb{E}[\pi \mid \text{win}] = q\mu_J - q\mathbb{E}[S] - C.$$

Substituting (★):

$$\mathbb{E}[\pi \mid \text{win}] = q\mu_J - q\left(\sigma_D\sqrt{\frac{2\tau_{\text{self}}}{\pi}} + \lambda\tau_{\text{self}}\mu_J\right) - C.$$

The strategy is profitable in expectation if $\mathbb{E}[\pi \mid \text{win}] > 0$, which (after dividing by q and rearranging) reduces to the inequality:

$$\mu_J > \sigma_D\sqrt{\frac{2\tau_{\text{self}}}{\pi}} + \lambda\tau_{\text{self}}\mu_J + \frac{C}{q}. \quad (\star\star)$$

The break-even latency τ^* is the unique positive root of the equation obtained by replacing the inequality in (★) with an equality:

$$\sigma_D\sqrt{\frac{2\tau^*}{\pi}} + \lambda\tau^*\mu_J + \frac{C}{q} = \mu_J. \quad (\dagger)$$

This is a quadratic in $\sqrt{\tau^*}$ and admits a closed-form solution. Let $u \equiv \sqrt{\tau^*}$, $a \equiv \lambda\mu_J$, $b \equiv \sigma_D\sqrt{2/\pi}$, $c \equiv C/q - \mu_J$ (note $c < 0$ when the strategy can be profitable at zero latency). Then:

$$au^2 + bu + c = 0, \quad u = \frac{-b + \sqrt{b^2 - 4ac}}{2a}, \quad \tau^* = u^2.$$

Asymptotic regimes. Two limits are useful for intuition.

- **Low intensity** ($\lambda \rightarrow 0$): the linear term in (†) vanishes, leaving $\sigma_D\sqrt{2\tau^*/\pi} = \mu_J - C/q$, so:

$$\tau^* \approx \frac{\pi}{2} \left(\frac{\mu_J - C/q}{\sigma_D} \right)^2.$$

Smaller diffusion σ_D or larger jump magnitude μ_J (relative to fixed cost) produce a larger break-even threshold.

- **High intensity** ($\lambda \rightarrow \infty$): the diffusion term becomes negligible relative to the linear term, leaving $\lambda\tau^*\mu_J \approx \mu_J - C/q$, so:

$$\tau^* \approx \frac{1}{\lambda} \left(1 - \frac{C}{q\mu_J} \right).$$

The threshold collapses as the inverse of event arrival rate.

Both expressions match intuition: the break-even latency tightens as the price process becomes faster (higher σ_D or higher λ), and loosens as the strategy's cost-to-jump ratio improves.

4.2 N-trader latency race

Suppose now there are N homogeneous arbitrageurs with i.i.d. latencies $\tau_1, \dots, \tau_N$ drawn from a distribution F . The trader is fastest with probability:

$$p_{\text{win}}^{(N)} = \mathbb{P}\left(\tau_{\text{self}} = \min_{i \in \{1, \dots, N\}} \tau_i\right) = \int_0^\infty (1 - F(\tau))^{N-1} f(\tau) d\tau.$$

For symmetric N -equal-latency competitors, $p_{\text{win}}^{(N)} = 1/N$. Expected profit per event therefore scales as:

$$\mathbb{E}[\pi^{(N)}] = \frac{1}{N}(q\mu_J - q\mathbb{E}[S] - C).$$

The marginal entrant’s effect is sharp: going from $N = 1$ to $N = 2$ halves expected profit; from $N = 3$ to $N = 4$ reduces by 25%; and so on. This is the standard Bertrand-in-speed result, modified by the slippage and cost terms inherited from the single-arbitrageur model.

When latencies are heterogeneous — as is empirical reality — the fastest participant captures a disproportionate share. If F is exponential with rate β , the order statistics of $\tau_{(1)}, \dots, \tau_{(N)}$ have known closed forms; the fastest’s expected profit share scales asymptotically as $\beta/(N\beta) = 1/N$ in the symmetric case, but for the empirically relevant log-normal latency distribution the fastest’s share grows more rapidly with the dispersion of F .

4.3 The role of broker-side defenses

The model so far assumes constant broker behavior. In practice, brokers deploy execution-side adjustments — collectively documented as “anti-arbitrage plugins”’ Fesenko, 2026a — which shift the slippage distribution rightward and the win-probability distribution leftward.

We model three concrete adjustments:

- **Asymmetric slippage** (a slippage engine biasing fills against the trader) introduces a multiplicative factor $\kappa_S \geq 1$ on $\mathbb{E}[S]$.
- **Last-look hold windows** introduce a probability $p_R \in [0, 1)$ that an otherwise-profitable order is rejected during the broker’s hold window, with the trader incurring a small cancel cost C_R .
- **Dynamic spread widening** raises C in a state-contingent way during the windows the trader is most likely to act.

Combining, expected profit per event becomes:

$$\mathbb{E}[\pi^{\text{def}}] = p_{\text{win}}[(1 - p_R)(q\mu_J - q\kappa_S\mathbb{E}[S] - C) - p_R C_R].$$

The break-even latency in the presence of defenses, τ_{def}^* , lies strictly above the undefended τ^* . Substituting the canonical defense parameters reported in Fesenko, 2026a ($\kappa_S = 1.5$, $p_R = 0.05$,

30% uplift on C during defense windows) yields a $\tau_{\text{def}}^*/\tau^*$ ratio of approximately 1.5 in the 2026 baseline calibration of Section 5.

5. Numerical Illustration

We calibrate the model to retail EUR/USD tick data for 2025–2026 using publicly available archives.

5.1 Calibration

Parameter	Symbol	Value (2026)	Source
Diffusion volatility	σ_D	0.012 pip $\cdot \text{s}^{-1/2}$	Tick-level estimation, EUR/USD
Event arrival rate	λ	12.4 s^{-1}	Microstructure events / second
Jump magnitude	μ_J	0.18 pip	Mean of $ \Delta P $ at events
Order size	q	1 standard lot	100,000 EUR notional
Per-trade cost	C	\$7 (≈ 0.7 pip)	0.5-pip spread + \$2 commission

The diffusion volatility σ_D is estimated from tick-level mid-quote returns over event-free windows; the event arrival rate λ is estimated as the per-second density of jumps exceeding three diffusion standard deviations on the same data; the jump magnitude μ_J is the empirical mean of the conditional $|\Delta P|$ in those events. Cost figures correspond to a typical retail aggregator account.

5.2 Break-even latency

Substituting into (†) and solving the quadratic in $u = \sqrt{\tau^*}$ yields:

$$\tau^* \approx 14 \text{ ms.}$$

This is the break-even latency on retail EUR/USD under typical liquidity conditions in 2026. The strategy is profitable in expectation only if $\tau_{\text{self}} < 14$ ms.

For comparison, calibrated to a 2010 baseline ($\sigma_D \approx 0.020$, $\lambda \approx 6$, $\mu_J \approx 0.32$, $C \approx 1.3$ pip), the same calculation gives $\tau^* \approx 65$ ms — illustrating the empirical compression of the latency edge over the past decade by approximately a factor of four.

5.3 Sensitivity

We compute τ^* under perturbations of each parameter, holding others at the 2026 baseline:

Perturbation	τ^* (ms)	Change
Baseline 2026	14	—
Low-volatility regime ($\sigma_D \times 0.5$)	28	+14
High-volatility regime ($\sigma_D \times 2$)	5	−9
News window ($\lambda \times 5$)	4	−10
Tighter spread ($C \times 0.5$)	22	+8
ECN-quality fees ($C \times 0.25$)	31	+17
Plus typical defenses ($\kappa_S = 1.5$, $p_R = 0.05$, $C \times 1.3$)	22	+8

Three observations follow.

First, τ^* is most sensitive to volatility — a doubling of σ_D collapses the break-even threshold by more than 60%. This explains why latency arbitrage is hardest precisely during the high-volatility events that generate the largest jumps: the larger jump magnitude raises the gross opportunity, but the larger diffusion volatility accelerates the slippage cost more than proportionally.

Second, fee structure matters substantially: an ECN-quality fee structure roughly doubles the survivable latency relative to a standard retail aggregator account. This is consistent with the empirical observation that institutional desks routinely operate at 30–50 ms latency budgets that would be unprofitable for retail.

Third, the addition of broker-side defenses shifts τ^* from 14 ms to approximately 22 ms. Many retail VPS configurations operate in the 15–25 ms range, meaning that the introduction of standard defenses can move a previously-marginally-profitable retail setup across the boundary into negative expected profit.

5.4 N-competitor race

Holding parameters at the 2026 baseline and varying the number of equal-latency competitors:

N	$p_{\text{win}}^{(N)}$	$\mathbb{E}[\pi^{(N)}]$ per event	Cumulative ratio vs. $N = 1$
1	1.00	\$11.00	1.00
2	0.50	\$5.50	0.50
3	0.33	\$3.67	0.33
4	0.25	\$2.75	0.25
8	0.13	\$1.38	0.13

The fourth equal competitor reduces individual expected profit per event by 75% relative to the monopoly case. Beyond $N = 4$ the residual profit is small enough that fixed infrastructure costs (VPS, market data) typically push individual participants below break-even.

This illustrates the classical race-to-the-bottom dynamic: the per-event profit of any individual participant declines hyperbolically in N , so absent entry barriers — such as heterogeneous infrastructure costs or broker-segmentation effects discussed in Section 6 — the long-run profit per latency arbitrageur tends to zero.

6. Discussion

6.1 Implications for retail strategies

The empirical 14 ms threshold under 2026 baseline conditions means that retail traders running on commercial VPS infrastructure — where typical ping to the broker matching engine is 5–40 ms depending on geography and provider quality — operate within a narrow band of profitability. The strategy is viable only with co-located or near-co-located VPS, well-aggregated quote feeds, and a counterparty (broker) that does not deploy active execution-side defenses.

The 22 ms post-defense threshold sharpens the practical conclusion: a retail arbitrageur on a 30 ms VPS connecting to a defensive broker operates at negative expected value. Detection of broker defenses — in the sense of Fesenko, 2026a — is a necessary preliminary to the strategy.

6.2 The race-to-the-bottom terminal value

Under our N -competitor model, expected profit declines hyperbolically in N , so absent entry barriers the long-run profit per participant approaches zero. Two forces preserve some positive equilibrium profit.

Heterogeneous infrastructure costs. Co-location and ECN-quality fee tiers are expensive; only a small number of participants invest. Heterogeneity in F_τ keeps the upper tail of the profit-share distribution thick and preserves a positive expected profit for the fastest participants.

Brokerage segmentation. The retail FX market is fragmented across hundreds of brokers with heterogeneous defenses, latencies, and fee structures. A skilled participant who selects brokers carefully effectively faces a smaller N at any given moment than the global headcount of retail latency arbitrageurs would suggest.

The combination preserves a positive equilibrium for well-positioned retail and small-institutional participants, although meaningfully smaller than the 2010 vintage.

6.3 Strategic alternatives

Practitioners increasingly migrate from pure latency arbitrage to related strategies whose profitability is less sensitive to the latency threshold.

News arbitrage (scheduled-event arbitrage) exploits the predictable λ -spike around macro releases. The model in Section 3 implies that during a news release, τ^* collapses (Section 5.3 shows a $5\times$ rise in λ moves τ^* from 14 ms to 4 ms). However, the *gross* opportunity per event also rises, because μ_J during macro releases is several times its event-free baseline. The result is that news arbitrage

favors traders with somewhat slower latency than pure tick arbitrage but faster decision logic and more sophisticated event interpretation.

Lock arbitrage (cross-broker hedge) opens offsetting positions across two brokers and harvests the slower side’s slippage; the strategy is latency-insensitive in the sense that both sides drift simultaneously, but capital-intensive (capital is committed on both sides) and constrained by margin and broker correlation.

Pair and triangular arbitrage exploit cross-pair pricing inefficiencies; not strictly latency-bound but benefit from low τ_{self} .

These alternatives are surveyed in the practitioner literature (Fesenko, 2026a, 2026b).

6.4 Limitations of the present model

Three simplifications limit the scope of the analytical results.

First, the model is single-pair. A multi-pair extension would let us quantify cross-pair correlation effects, which matter when the trader’s signal feed aggregates across many pairs simultaneously.

Second, the broker-defense model in Section 4.3 is reduced-form. A structural model in which the broker chooses defense intensity to maximize expected revenue (taking the trader’s strategy as given) would be illuminating; such a model would predict the equilibrium defense intensity as a function of broker order-flow characteristics and is left to future work.

Third, the empirical calibration in Section 5 is illustrative. A formal estimation on a multi-broker tick panel — ideally with the broker-defense indicators of Fesenko, 2026a used as controls — would tighten the confidence bands on τ^* substantially. A natural empirical extension is to estimate τ^* separately by broker tier and to test the implied cross-sectional dispersion against the $1/N$ prediction of Section 4.2.

7. Conclusion

We have presented a closed-form decomposition of the expected per-trade profit of a retail latency-arbitrage strategy in foreign exchange and derived a break-even latency τ^* as a function of (i) volatility, (ii) signal arrival rate, and (iii) per-trade fixed cost. Calibrating to retail EUR/USD tick data for 2025–2026, the break-even threshold sits at approximately 14 ms — roughly a quarter of its 2010 value, consistent with the widely reported empirical compression of the latency edge in retail FX over the past decade.

The model also makes explicit two effects. The multiplicative effect of broker-side execution defenses raises the practical break-even latency by roughly 50%, moving many retail VPS configurations across the boundary into negative expected profit. The $1/N$ decay of profit per participant in a multi-competitor race pushes the long-run profit per arbitrageur toward zero in the absence of infrastructure heterogeneity or broker segmentation. Together these effects predict — and empirically

support — the continued migration of retail FX participants toward news-arbitrage, lock-arbitrage, and pair-arbitrage strategies whose profitability depends less critically on the latency margin.

Open research directions include a structural model of broker defense intensity, a multi-broker tick-panel estimation of τ^* with defense-indicator controls, and a multi-pair extension that quantifies cross-pair correlation effects. The practitioner literature accumulating around the broker-defense problem (Fesenko, 2026a) suggests that such extensions are increasingly tractable as more of the underlying execution data becomes available to retail and small-institutional researchers.

References

- Aquilina, M., Budish, E., & O'Neill, P. (2022). Quantifying the high-frequency trading “arms race”. *Quarterly Journal of Economics*, 137(1), 493–564.
- Brogaard, J., Hendershott, T., & Riordan, R. (2014). High-frequency trading and price discovery. *Review of Financial Studies*, 27(8), 2267–2306.
- Budish, E., Cramton, P., & Shim, J. (2015). The high-frequency trading arms race: Frequent batch auctions as a market design response. *Quarterly Journal of Economics*, 130(4), 1547–1621.
- Easley, D., López de Prado, M. M., & O'Hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. *Review of Financial Studies*, 25(5), 1457–1493.
- Fesenko, B. (2026a). Seven anti-arbitrage plugins brokers use — named, with detection mechanics. *BJF Trading Group, working paper*. <https://bjftradinggroup.com/anti-arbitrage-plugins/>
- Fesenko, B. (2026b). FIX session debugging: Top ten broker-dialect quirks that break your client. *BJF Trading Group, technical note*. <https://bjftradinggroup.com/fix-session-debugging/>
- Foucault, T., Hombert, J., & Roşu, I. (2016). News trading and speed. *Journal of Finance*, 71(1), 335–382.
- Foucault, T., Kadan, O., & Kandel, E. (2013). Liquidity cycles and make/take fees in electronic markets. *Journal of Finance*, 68(1), 299–341.
- Hasbrouck, J. (2018). High-frequency quoting: Short-term volatility in bids and offers. *Journal of Financial and Quantitative Analysis*, 53(2), 613–641.
- Hendershott, T., Jones, C. M., & Menkveld, A. J. (2011). Does algorithmic trading improve liquidity? *Journal of Finance*, 66(1), 1–33.
- Menkveld, A. J. (2013). High frequency trading and the new market makers. *Journal of Financial Markets*, 16(4), 712–740.
- Menkveld, A. J. (2016). The economics of high-frequency trading: Taking stock. *Annual Review of Financial Economics*, 8, 1–24.
- O'Hara, M. (2015). High frequency market microstructure. *Journal of Financial Economics*, 116(2), 257–270.

Stoll, H. R. (1989). Inferring the components of the bid-ask spread: Theory and empirical tests. *Journal of Finance*, 44(1), 115–134.

Stoll, H. R. (2000). Friction. *Journal of Finance*, 55(4), 1479–1514.