Mathematics and Statistics of Algorithmic Trading[1]

University of Oxford

Álvaro Cartea

alvaro.cartea@maths.ox.ac.uk
http://www.cartea.net

February 25, 2016

[1]Notes based on textbook "Algorithmic and High-Frequency Trading" with Sebastian Jaimungal and Jose Penalva.

Optimal Execution

- How do we trade in or out of a large position when Order-Flow affects prices?
- Optimal Execution with Temporary Impact and Inventory Control

Permanent Impact

We assume a linear relationship between net order-flow and changes in the midprice, thus for every trading day we perform the regression

$$\Delta S_n = b\,\mu_n + \varepsilon_n \tag{1}$$

where $\Delta S_n = S_{n\tau} - S_{(n-1)\tau}$ is the change in the midprice, μ_n is net order-flow defined as the difference between the volume of buy and sell MOs during the time interval $[(n-1)\tau,\, n\tau]$, and ε_n is the error term (assumed normal). In the empirical analysis we choose $\tau = 5\,\text{min}$.

Temporary Impact

- Assume that temporary price impact is linear in the rate of trading so the difference between the execution price that the investor receives and the midprice is $k\,\nu$,
- To do this we take a snapshot of the LOB each second,
- Determine the price per share for various volumes (by walking through the LOB),
- Compute the difference between the price per share and the best quote at that time,
- Perform a linear regression.

Parameters

Table: Permanent and temporary price impact parameters for Nasdaq stocks, average volume of MOs, average midprice, σ volatility (hourly) of arithmetic price changes, mean arrival (hourly) of MOs $\lambda^{\pm}$, and average volume of MOs $\mathbb{E}[\eta^{\pm}]$. Data are from Nasdaq 2013.

	FARO		SMH		NTAP	
	mean	stdev	mean	stdev	mean	stdev
ADV	23,914	14,954	233,609	148,580	1,209,628	642,376
midprice	40.55	6.71	37.90	2.44	38.33	3.20
σ	0.151	0.077	0.067	0.039	0.078	0.045
b	1.41E-04	9.61E-05	5.45E-06	4.20E-06	5.93E-06	2.31E-06
k	1.86E-04	2.56E-04	8.49E-07	8.22E-07	3.09E-06	1.75E-06
b/k	1.02	0.83	7.43	6.24	2.04	0.77
λ^+	16.81	9.45	47.29	28.13	300.52	144.48
$\mathbb{E}[\eta^+]$	103.56	21.16	377.05	118.05	308.45	53.09
λ^-	17.62	10.69	46.37	27.62	293.83	136.13
$\mathbb{E}[\eta^-]$	104.00	21.79	381.70	126.74	312.81	49.86

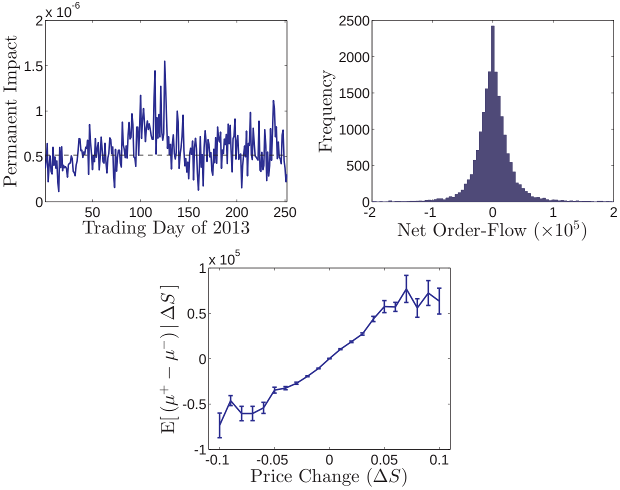


Figure: Order-Flow and effect on the drift of midprice of INTC. The first picture shows

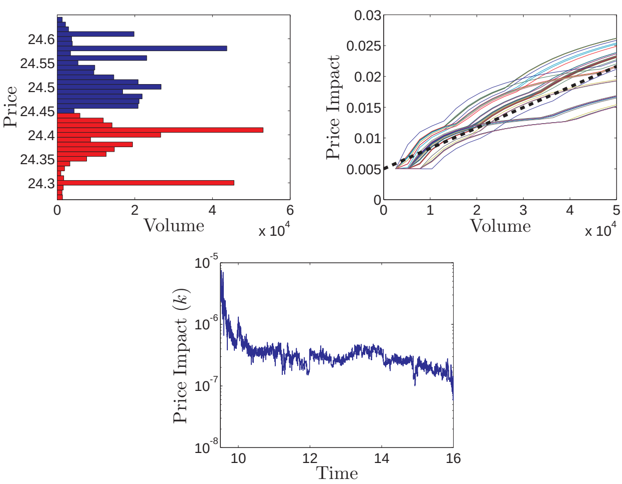


Figure: An illustration of how the temporary impact may be estimated from snapshots of the LOB using INTC on Nov 1, 2013. The first panel is at 11:00am, the second from 11:00am to 11:01am and the third contains the entire day.

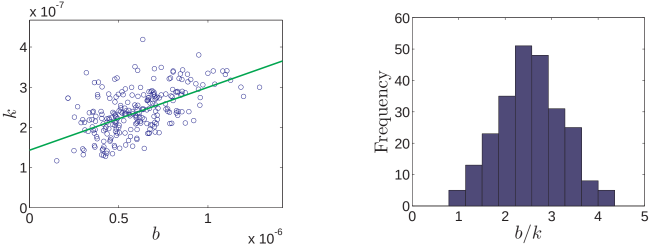


Figure: Price Impact INTC using daily observations for 2013.

The Model

- Inventory (investor is liquidating)

$$dQ_t^\nu = -\nu_t \, dt\,, \qquad Q_0^\nu = \mathfrak{N}\,, \tag{2}$$

- Midprice

$$dS_t^\nu = b\,(\mu_t - \nu_t)\,dt + \sigma\,dW_t\,, \tag{3}$$

  where $\mu_t = \mu_t^+ - \mu-_t$ is net order-flow of other market participants.
- Execution price

$$\hat{S}_t^\nu = S_t^\nu - \left(\tfrac{1}{2}\Delta + k\,\nu_t)\right)\,, \tag{4}$$

- Cash

$$dX_t^v = \hat{S}_t^\nu\,\nu_t\,dt\,, \qquad X_0^v = x\,. \tag{5}$$

Value Function

$$H^{\nu}() = \mathbb{E}_{t,x,S,\boldsymbol{\mu},q}\left[X_{\tau^\nu} + Q^{\nu}_{\tau^\nu}\;\left(S^{\nu}_{\tau^\nu} - \tfrac{1}{2}\Delta - \alpha Q^{\nu}_{\tau^\nu}\right) - \phi\int_t^{\tau^\nu} (Q^{\nu}_u)^2\,du\right] \tag{6}$$

where $\tau^\nu = T \wedge \inf\{t\,:\,Q^\nu_t = 0\}$, and $\boldsymbol{\mu} = \{\mu^+,\mu^-\}$. Her value function is

$$H(t,x,S,\boldsymbol{\mu},q) = \sup_{\nu\in\mathcal{A}} H^\nu(t,x,S,\boldsymbol{\mu},q)\,, \tag{7}$$

The DPP suggests that $H(t,x,S,\mu,q)$ satisfies

$$0 = \left(\partial_t + \tfrac{1}{2}\sigma^2\,\partial_{SS}\right)H + \mathcal{L}^{\boldsymbol{\mu}}H - \phi\,q^2 + \sup_{\nu}\left\{\Big(\nu\,(S - \tfrac{1}{2}\Delta - k\,\nu)\,\partial_x + b\,(\mu-\nu)\,\partial_S - \nu\,\partial_q\Big)H\right\} \tag{8}$$

for $q>0$, subject to the terminal condition.

$$H(T,x,S,\boldsymbol{\mu},q) = x + q\;\left(S - \tfrac{1}{2}\Delta\right) - \alpha\,q^2\,.$$

Solving the DPE.

The DPE (8) admits the solution

$$H(t,x,S,\boldsymbol{\mu},q) = x + q\,\left(S - \tfrac{1}{2}\Delta\right) + h_0(t,\boldsymbol{\mu}) + q\,h_1(t,\boldsymbol{\mu}) + q^2\,h_2(t)\,, \qquad q > 0\,,$$

where

$$h_2(t) = \sqrt{k\,\phi}\,\frac{1+\zeta\,e^{2\gamma(T-t)}}{1-\zeta\,e^{2\gamma(T-t)}} - \tfrac{1}{2}\,b\,, \tag{9a}$$

$$h_1(t,\boldsymbol{\mu}) = b\int_t^T \left(\frac{e^{-\gamma(T-u)} - \zeta e^{\gamma(T-u)}}{e^{-\gamma(T-t)} - \zeta e^{\gamma(T-t)}}\right)\mathbb{E}_{t,\boldsymbol{\mu}}\left[\,\mu_u\,\right]\,du\,, \tag{9b}$$

$$h_0(t,\boldsymbol{\mu}) = \frac{1}{4k}\int_t^T \mathbb{E}_{t,\boldsymbol{\mu}}\left[\,h_1^2(t,\boldsymbol{\mu}_u)\,\right]\,du\,, \tag{9c}$$

with the constants γ and ζ:

$$\gamma = \sqrt{\frac{\phi}{k}}, \quad \text{and} \quad \zeta = \frac{\alpha - \frac{1}{2}b + \sqrt{k\,\phi}}{\alpha - \frac{1}{2}b - \sqrt{k\,\phi}}\,.$$

Proof

To solve (8) we make the ansatz

$$H(t,x,S,\boldsymbol{\mu},q) = x + q\,\left(S - \tfrac{1}{2}\Delta\right) + h(t,\boldsymbol{\mu},q)\,,$$

and upon substitution of the ansatz in the DPE $h(t,\boldsymbol{\mu},q)$ satisfies

$$\partial_t h + \mathcal{L}^{\boldsymbol{\mu}} h + b\,\mu\,q - \phi\,q^2 + \sup_{\nu}\left\{-k\,\nu^2 - (b\,q + \partial_q h)\,\nu\right\} = 0\,, \qquad q > 0\,,$$

subject to $h(T,\boldsymbol{\mu},q) = -\alpha\,q^2$, so

$$\nu^* = -\frac{1}{2\,k}\,(b\,q + \partial_q h)\;. \tag{10}$$

Upon substitution back into the DPE we find that h satisfies

$$(\partial_t + \mathcal{L}^{\boldsymbol{\mu}})\,h + b\,\mu\,q - \phi\,q^2 + \frac{1}{4\,k}\,(b\,q + \partial_q h)^2 = 0\,, \qquad q > 0\,. \tag{11}$$

Proof (cont)

Due to the existence of linear and quadratic terms in q in (11), and its terminal conditions, we expect $h(t, \boldsymbol{\mu}, q)$ to be a quadratic form in q, and we assume the ansatz

$$h(t, \boldsymbol{\mu}, q) = h_0(t, \boldsymbol{\mu}) + q\, h_1(t, \boldsymbol{\mu}) + q^2\, h_2(t, \boldsymbol{\mu})\,.$$

Inserting this into (11) and collecting like terms in q leads to

$$\left(\partial_t + \mathcal{L}^{\boldsymbol{\mu}}\right) h_0 + \frac{1}{4k}\, h_1^2 = 0\,, \tag{12a}$$

$$\left(\partial_t + \mathcal{L}^{\boldsymbol{\mu}}\right) h_1 + b\,\mu + \frac{1}{2k}\, h_1\, \left(b + 2h_2\right) = 0\,, \tag{12b}$$

$$\left(\partial_t + \mathcal{L}^{\boldsymbol{\mu}}\right) h_2 - \phi + \frac{1}{4k}\, \left(b + 2h_2\right)^2 = 0\,, \tag{12c}$$

subject to the terminal conditions

$$h_0(T, \boldsymbol{\mu}) = 0, \quad h_1(T, \boldsymbol{\mu}) = 0, \quad h_2(T, \boldsymbol{\mu}) = -\alpha\,.$$

Proof (cont)

To solve for h_2 we note that since Equation (12c) for h_2 contains no source terms in μ and its terminal condition is independent of μ, the solution must be independent of μ, i.e. h_2 is a function only of time. In this case, (12c) is an ODE of Riccati type and can be solved explicitly:

$$h_2(t, \boldsymbol{\mu}) = \chi(t) - \tfrac{1}{2}\, b, \quad \text{where} \quad \chi(t) = \sqrt{k\,\phi}\,\frac{1 + \zeta\, e^{2\gamma(T-t)}}{1 - \zeta\, e^{2\gamma(T-t)}}\,,$$

with the constants γ and ζ:

$$\gamma = \sqrt{\frac{\phi}{k}}, \quad \text{and} \quad \zeta = \frac{\alpha - \frac{1}{2}\,b + \sqrt{k\,\phi}}{\alpha - \frac{1}{2}\,b - \sqrt{k\,\phi}}\,.$$

Proof (cont)

Now we turn to solving (12b) which is a linear PIDE for h_1 where $h_2 + \frac{1}{2}b$ acts as an effective discount rate and $b\,\mu$ is a source term. The general solution of such an equation can be represented using the Feynman-Kac theorem. Thus we write

$$h_1(t,\boldsymbol{\mu}) = b\,\mathbb{E}_{t,\boldsymbol{\mu}}\left[\int_t^T \exp\left\{\frac{1}{k}\int_t^u \left(h_2(s) + \tfrac{1}{2}b\right)\,ds\right\}\mu_u\,du\right]$$

which can be simplified to

$$h_1(t,\boldsymbol{\mu}) = b\int_t^T \left(\frac{e^{-\gamma(T-u)} - \zeta\,e^{\gamma(T-u)}}{e^{-\gamma(T-t)} - \zeta\,e^{\gamma(T-t)}}\right)\mathbb{E}_{t,\boldsymbol{\mu}}\left[\,\mu_u\,\right]du\,. \tag{13}$$

Finally, we can solve for $h_0(t,\boldsymbol{\mu})$ by again noticing it is a linear PDE with non-linear source term and a straight forward application of Feynman-Kac, and interchanging integration and expectation, we obtain (9c). $\square$

Limiting cases

All shares must be liquidated, $\alpha \to \infty$, and the optimal trading speed simplifies to

$$\lim_{\alpha\to\infty} \nu_t^* = \gamma \frac{\cosh(\gamma\,(T-t))}{\sinh(\gamma\,(T-t))}\, Q_t^{\nu^*} - \frac{b}{2k} \int_t^T \frac{\sinh(\gamma(T-u))}{\sinh(\gamma(T-t))}\, \mathbb{E}\left[\, \mu_u \,|\mathcal{F}_t^{\boldsymbol{\mu}}\,\right] du\,. \tag{14}$$

Another interesting limiting case is if the above limit is followed by letting $\phi \to 0$:

$$\lim_{\phi\to 0} \lim_{\alpha\to\infty} \nu_t^* = \frac{1}{(T-t)}\, Q_t^{\nu^*} - \frac{b}{2k} \int_t^T \frac{(T-u)}{(T-t)}\, \mathbb{E}\left[\, \mu_u \,|\mathcal{F}_t^{\boldsymbol{\mu}}\right] du\,. \tag{15}$$

Model for Order-Flow

Order-flow $\mu_t^{\pm}$ satisfy the SDEs

$$d\mu_t^{\pm} = -\kappa\,\mu_t^{\pm}\,dt + \eta^{\pm}_{1+L^{\pm}_{t-}}\,dL_t^{\pm}\,, \tag{16}$$

where $L_t^{\pm}$ are independent Poisson processes with equal intensity λ, $\{\eta_1^{\pm}, \eta_2^{\pm}, \dots\}$.
The solutions to (16), for $s > t$, are

$$\mu_s^{\pm} = e^{-\kappa^{\pm}(s-t)}\,\mu_t^{\pm} + \int_t^s e^{-\kappa^{\pm}(s-u)}\,\eta^{\pm}_{1+N^{\pm}_{u^-}}\,dL_u^{\pm}\,,$$

so that

$$\mathbb{E}[\,\mu_s^{\pm}\,|\,\mathcal{F}_t^{\mu}\,] = e^{-\kappa^{\pm}(s-t)}\,\left(\mu_t^{\pm} - \psi^{\pm}\right) + \psi^{\pm}\,,$$

where

$$\psi^{\pm} = \tfrac{1}{\kappa^{\pm}}\,\lambda^{\pm}\,\mathbb{E}[\eta^{\pm}]\,.$$

Therefore, under this particular model for order-flow, we follow the optimal trading strategy

$$\begin{aligned}\lim_{\alpha\to\infty}\nu_t^* &= \gamma\,\frac{\cosh\left(\gamma(T-t)\right)}{\sinh\left(\gamma(T-t)\right)}\,Q_t^{\nu^*}\\ &-\frac{b}{2\,k}\left[\,\ell_1^+(t)\left(\mu_t^+-\psi^+\right)-\ell_1^-(t)\left(\mu_t^--\psi^-\right)+\ell_0(t)\left(\psi^+-\psi^-\right)\,\right]\end{aligned}$$

where

$$\ell_0(t)=\frac{1}{\gamma}\,\frac{\cosh\left(\gamma(T-t)\right)-1}{\sinh\left(\gamma(T-t)\right)}\,,$$

and

$$\ell_1^{\pm}(t)=\frac{1}{2}\,\left(\frac{e^{\gamma(T-t)}-e^{-\kappa^{\pm}(T-t)}}{\kappa^{\pm}+\gamma}-\frac{e^{-\gamma(T-t)}-e^{-\kappa^{\pm}(T-t)}}{\kappa^{\pm}-\gamma}\right)\Bigg/\ \sinh\left(\gamma(T-t)\right)$$

Sims

- We assume that the trading horizon is $T = 1\,\text{hour}$,
- the running inventory penalty parameter is $\phi = 10 \times k$,
- the liquidation target $\mathfrak{N}$ shares is set to 1% of the expected traded volume over the trading window (including the investor's trades),
- run $10,000$ simulations,
- the strategy trades at the same frequency as that of the arrival of market sell orders from all other agents

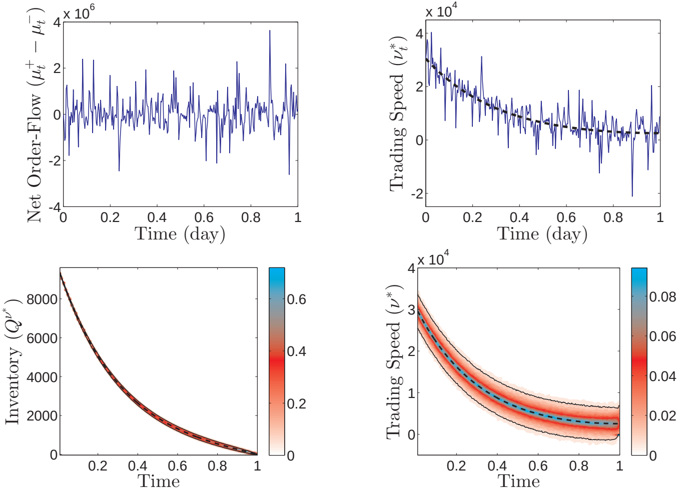
Use INTC with parameters

$$k = 2.50 \times 10^{-7}\,, \quad b = 6.15 \times 10^{-}\,, \quad \mathfrak{N} = 9,612\,,$$

$$\lambda^{\pm} = 328\,, \quad \kappa = 360\,, \quad \eta \sim \mathit{Exp}(1,453 \times \kappa)\,,$$

$$\sigma = 0.01\,, \quad S_0 = 23.04\,,$$

and $\eta \sim \mathit{Exp}(\eta_0)$ denotes the exponential distribution with mean size $\mathbb{E}[\eta] = \eta_0$,


x 10^6
Net Order-Flow $(\mu_t^+ - \mu_t^-)$
Time (day)
x 10^4
Trading Speed (ν_t^*)
Time (day)
Inventory (Q^{ν^*})
Time
x 10^4
Trading Speed (ν^*)
Time

Histogram of the financial performance of the strategy relative to Almgren-Chriss. This performance is measured in basis points using

$$\frac{X_T^{\nu^*} - X_T^{AC}}{X_T^{AC}} \times 10^4 \tag{17}$$

where X_T^{AC} is the cash obtained from running the Almgren-Chriss strategy with the same level of ϕ.

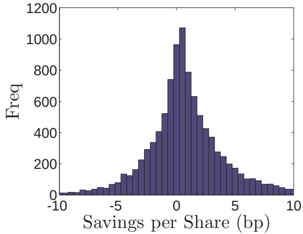


Figure: The savings per share (INTC) measured in basis points, $(X_T^{\nu^*} - X_T^{AC})/X_T^{AC} \times 10^4$, from following the optimal strategy relative to Almgren-Chriss.

Table: Simulation performance relative to AC

		ν^*				
		FARO	SMH	NTAP	ORCL	INTC
	mean	529	602	0.78	0.64	0.93
	stdev	739	708	3.71	3.02	3.78
quantile	5%	-438	5	-4.70	-3.83	-4.48
quantile	25%	151	157	-0.98	-0.82	-0.87
quantile	50%	396	428	0.50	0.44	0.55
quantile	75%	835	895	2.23	1.85	2.45
quantile	95%	1862	1957	6.98	5.58	7.51
	$\%t\, \nu_t^* < 0$	12.3%	17.2%	2.7%	2.6%	3.4%
	$\%X_T < X_T^{AC}$	10.2%	4.9%	38.9%	39.2%	38.5%

Table: Simulation performance relative to AC

		$\max(\nu^*, 0)$				
		FARO	SMH	NTAP	ORCL	INTC
	mean	615	670	1.12	0.91	1.43
	stdev	648	654	3.80	3.14	3.97
quantile	5%	-1	42	-4.10	-3.46	-3.79
quantile	25%	179	203	-0.79	-0.67	-0.58
quantile	50%	434	470	0.61	0.52	0.76
quantile	75%	868	942	2.52	2.08	2.89
quantile	95%	1887	1952	7.85	6.41	8.57
$\%t\, \nu_t^* < 0$		0.0%	0.0%	0.0%	0.0%	0.0%
$\%X_T < X_T^{AC}$		5.0%	0.1%	36.5%	37.0%	34.7%