Limit order books: Tractable SPDE models

Marvin S. Müller (ETHZ)

joint work with R. Cont (Oxford)

Innovative Research in Mathematical Finance

celebration of 70 years of Yuri Kabanov

CIRM, Luminy, 7th September 2018

The Beginning

Bachelier (1900): BM as stock price model at Paris Bourse:

$$\mathrm{d}s_t = \mu \,\mathrm{d}t + \sigma \,\mathrm{d}B_t, \quad t \geq 0,$$

B BM, $\mu \in \mathbb{R}$, $\sigma > 0$.

The Beginning

Bachelier (1900): BM as stock price model at Paris Bourse:

$$ds_t = \mu \, dt + \sigma \, dB_t, \quad t \geq 0,$$

B BM, $\mu \in \mathbb{R}$, $\sigma > 0$.

- 30y later: Rigorous framework due to Kolmogorov
- Several extensions: Black-Scholes-Merton, local-vol, stochastic-vol etc.

The Beginning

Bachelier (1900): BM as stock price model at Paris Bourse:

$$\mathrm{d}s_t = \mu \,\mathrm{d}t + \sigma \,\mathrm{d}B_t, \quad t \geq 0,$$

B BM, $\mu \in \mathbb{R}$, $\sigma > 0$.

- 30y later: Rigorous framework due to Kolmogorov
- Several extensions: Black-Scholes-Merton, local-vol, stochastic-vol etc.
- 77y later: Toronto Stock Exchange introduced CATS (Computer Assisted Trading System)

Exchanges and LOBs

Orders are submitted and electronically stored /matched.

- Limit Order: buy / sell quantity of q orders at price level p or better
- Market Order: buy / sell quantity q at best market price, executed directly against matching LOs

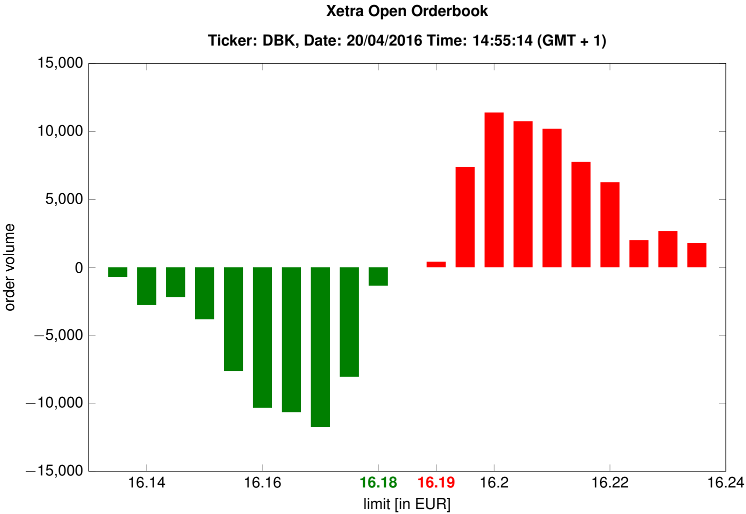
Exchanges and LOBs

Orders are submitted and electronically stored /matched.

- Limit Order: buy / sell quantity of q orders at price level p or better
- Market Order: buy / sell quantity q at best market price, executed directly against matching LOs

- Execution of market orders decreased from sec or min to less than msec
- Much more information available than bid-ask quotes
- In 2002: more than 60 % of world largest stock exchanges are fully order driven, 80 % at least partially. (Jain (2005))

Limit Order Books



High-Dimensional Modeling

Mid price: $s := {}^{(p_{\text{bid}} + p_{\text{ask}})}/_{2}$

LOB Model: $v_t(p)$ density of LOB, centered: $u_t(p) := v_t(p + s_t)$

Observations and Assumptions

- HFT: $>$ 1000 orders per 10sec on average for some US stocks (Cont et al 2011)
- On average, orders arrive at prices p

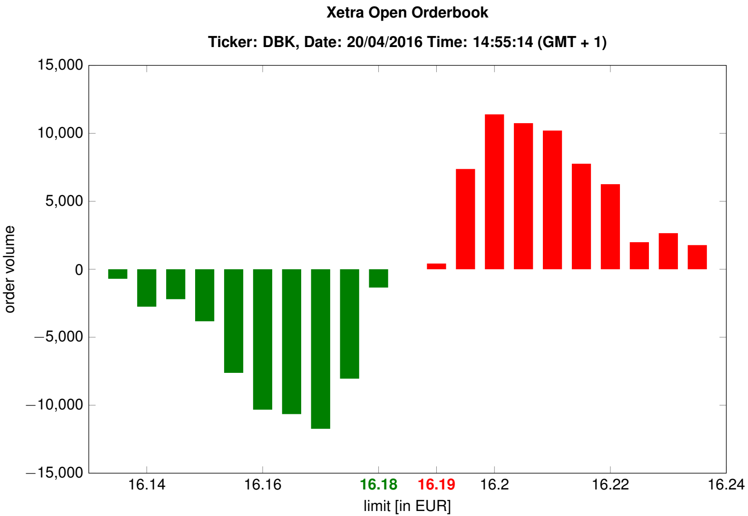
$$\sim a(b + |p - s_t|)^{-1-\mu}, \qquad \mu \in [0.6, 1.5]$$

(Bouchaud et al (2002), Zovko, Farmer (2006))

$\longrightarrow$ price-time-continuous model,

$\longrightarrow$ no spread, $p_{\text{bid}} = p_{\text{ask}} = s$.

Limit Order Books: Discrete Reality



Limit Order Books: Density $v_t(p)$

Space time continuous apprpoxixmation:

Macroscopic Order Book Dynamics I

Fixed frame: $u_t(p) := v_t(p + s_t)$:

1. Small order readjustments of HF-traders: rate η
2. Tendency to shift orders in direction to bid/ask (HF-traders): rate β
3. Net impact rate for order volume, of LF and HF-traders: α
4. LF-net impact due to exogeneous information: $f(p)$
5. HF-trader impact on volume: $\mathrm{d}X_t$

Order book density

With $\mathrm{d}p,\ \mathrm{d}t \to 0$, we impose for the centered order book density

$$\mathrm{d}u_t(p) = \underbrace{[\eta_a \Delta u_t(p) + \beta_a \nabla u_t(p) + \alpha_a u_t(p) + f_a(p)]}_{=A_a u_t + f_a} \mathrm{d}t + u_{t-}(p)\ \mathrm{d}X_t^a.$$

for $p > 0$ (< 0 resp.) where X^a is a $\mathbb{R}$-valued semimartingale.

Linear Models

On an abstract level,

$$\begin{cases} \mathrm{d}u_t(p) = A_a u_t(p)\,\mathrm{d}t + u_{t-}(p)\,\mathrm{d}X_t^a, & p \in (0, L), \\ \mathrm{d}u_t(p) = A_b u_t(p)\,\mathrm{d}t + u_{t-}(p)\,\mathrm{d}X_t^b, & p \in (-L, 0). \end{cases} \tag{1}$$

X^a, X^b are real semimartingales w/ jumps > -1 and A_a, A_b are densely defined and closed linear maps on $L^2(-L, L)$

Linear Models

On an abstract level,

$$\begin{cases} \mathrm{d}u_t(p) = A_a u_t(p)\,\mathrm{d}t + u_{t-}(p)\,\mathrm{d}X_t^a, & p \in (0, L), \\ \mathrm{d}u_t(p) = A_b u_t(p)\,\mathrm{d}t + u_{t-}(p)\,\mathrm{d}X_t^b, & p \in (-L, 0). \end{cases} \tag{1}$$

X^a, X^b are real semimartingales w/ jumps > -1 and A_a, A_b are densely defined and closed linear maps on $L^2(-L, L)$

Theorem

$g\colon [0,\infty) \to L^2(-L,L)$ solves (1) *for $X^{a/b} \equiv 0$, iff*

$$u_t(p) := g_t(p)\left[\mathcal{E}_t(X^b)\mathbf{1}_{(-L,0)}(p) + \mathcal{E}_t(X^a)\mathbf{1}_{(0,L)}(p)\right], \quad p \in (-L, L),$$

solves (1) *(in analytically weak sense).*

Examples

Ex 1: Level-1 Models

$L := \mathrm{d}p$, best bid and ask queues are modeled by positive semimartingales Z^a, Z^b. Then X^a and X^b are such that $Z^\star = \mathcal{E}(X^\star)$, $\star \in \{a, b\}$.

Examples

Ex 1: Level-1 Models

$L := \mathrm{d}p$, best bid and ask queues are modeled by positive semimartingales Z^a, Z^b. Then X^a and X^b are such that $Z^\star = \mathcal{E}(X^\star)$, $\star \in \{a, b\}$.

Ex 2: Intra-Book dynamics

$L > 0$ very large (e.g. $L = 1000$), A_a and A_b model intra-book dynamics, as above:

$$A_a := \eta_a \Delta + \beta_a \nabla + \alpha_a, \quad A_b := \eta_b \Delta - \beta_b \nabla + \alpha_b$$

Spectral Analysis

Imposing Dirichlet bdry cond., the eigenvalues and eigenfcts of $A := \eta\Delta + \beta\nabla + \alpha$ on $(0, L)$ are

$$\nu_k := \alpha - k^2 \frac{\eta\pi^2}{L^2} - \frac{\beta^2}{4\eta}, \quad e_k(p) := e^{-\frac{\beta}{2\eta}p} \sin\left(\frac{k\pi}{L}p\right), k \in \mathbb{N},$$

Observation

- Solution of $\mathrm{d}u_t = Au_t\,\mathrm{d}t + u_{t-}\,\mathrm{d}X_t$ is

$$u_t(p) = \mathcal{E}_t(X) \sum_{k=1}^{\infty} e^{\nu_k t} \left(\int_0^L u_0(x) e_k(x) e^{\frac{\beta}{\eta}x}\,\mathrm{d}x\right) e_k(p) \qquad (2)$$

- The only positive eigenfct is e_1 for the principlec eigenvalue $\nu := \nu_1$.

$$\begin{cases} \mathrm{d}u_t(p) = \left[\eta^a \Delta u_t(p) + \beta^a \nabla u_t(p) + \alpha^a u_t(p)\right] \mathrm{d}t + \sigma^a u_t(p)\, \mathrm{d}W_t^a, & p \in (0, L), \\ \mathrm{d}u_t(p) = \left[\eta^b \Delta u_t(p) - \beta^b \nabla u_t(p) + \alpha^b u_t(p)\right] \mathrm{d}t + \sigma^b u_t(p)\, \mathrm{d}W_t^b, & p \in (-L, 0), \\ u_t(0+) = u_t(0-) = u(-L) = u(L) = 0, \\ \quad u_t(p) > 0, \quad p \in (0, L),\ u_t(p) < 0, \quad p \in (-L, 0), \quad t > 0, \end{cases}$$

From parametrization theorem: $u_t(p)$ explicitly computable

$$\begin{cases} \mathrm{d}u_t(p) = \left[\eta^a \Delta u_t(p) + \beta^a \nabla u_t(p) + \alpha^a u_t(p)\right] \mathrm{d}t + \sigma^a u_t(p)\, \mathrm{d}W_t^a, & p \in (0, L), \\ \mathrm{d}u_t(p) = \left[\eta^b \Delta u_t(p) - \beta^b \nabla u_t(p) + \alpha^b u_t(p)\right] \mathrm{d}t + \sigma^b u_t(p)\, \mathrm{d}W_t^b, & p \in (-L, 0), \\ u_t(0+) = u_t(0-) = u(-L) = u(L) = 0, \\ \quad u_t(p) > 0, \quad p \in (0, L),\; u_t(p) < 0, \quad p \in (-L, 0), \quad t > 0, \end{cases}$$

From parametrization theorem: $u_t(p)$ explicitly computable , and

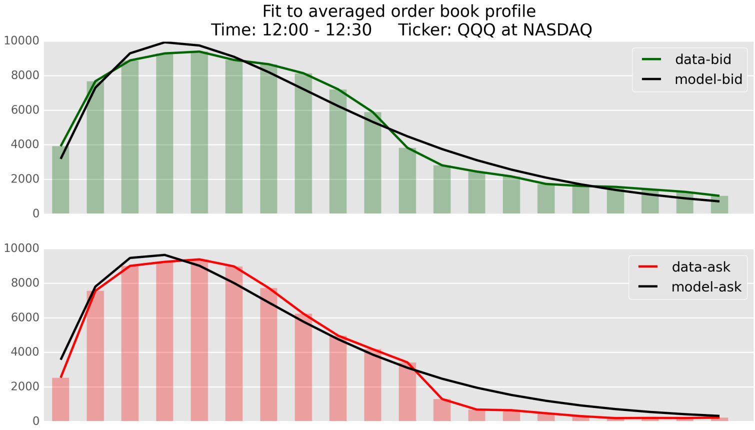
Corollary

For $u_0(p) = h(p) := \sin(\frac{\pi}{L} p) \exp\left(\pm \frac{\beta^{b/a}}{2\eta_{b/a}} p\right)$ the unique solution is

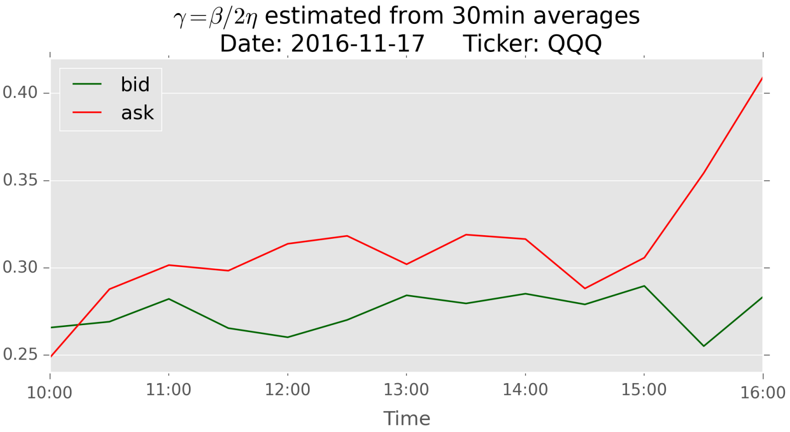
$$u_t(p) = h(p) \left(Y_t^b \mathbf{1}_{(-L,0)}(p) + Y_t^a \mathbf{1}_{(0,L)}(p) \right),$$

where $\mathrm{d}Y_t^{a/b} = \nu_{a/b} Y_t^{a/b}\, \mathrm{d}t + \sigma_{a/b} Y_t^{a/b}\, \mathrm{d}W_t^{a/b}$

Fit to data:


Fit to averaged order book profile
Time: 12:00 - 12:30 Ticker: QQQ at NASDAQ
10000
8000
6000
4000
2000
0
data-bid
model-bid
10000
8000
6000
4000
2000
0
data-ask
model-ask

Fit to data:


$\gamma=\beta/2\eta$ estimated from 30min averages
Date: 2016-11-17 Ticker: QQQ
bid
ask
0.40
0.35
0.30
0.25
10:00
11:00
12:00
13:00
14:00
15:00
16:00
Time

Linear Inhomogeneous Models II

Consider now the one-sided problems,

$$\mathrm{d}u_t(p) = [Au_t(p) + \lambda f(p)]\,\mathrm{d}t + \sigma u_t(p)\,\mathrm{d}W_t, \qquad u_0(p) = z_0 f(p), \tag{3}$$

$p \in (0, L)$, with initial data $z_0 \in \mathbb{R}$.

Theorem

- *If f is an eigenfct for A with eigenvalue $-\nu$, then $u_t(p) = f(p)Z_t$, where $Z_0 = z_0$ and*

$$\mathrm{d}Z_t = [\lambda - \nu Z_t]\,\mathrm{d}t + \sigma Z_t\,\mathrm{d}W_t.$$

- *If f is not an eigenfct, then no “reasonable” parametrization!*

Price Prediction

- Empirical observation (Cont et al. '13):

$$\mathrm{d}s_t^{b/a} \approx \pm\delta \frac{\mathrm{OF}_{b/a}(t)}{D_{b/a}(t)}, \qquad D \ldots \text{depths}, \quad \mathrm{OF} \ldots \text{order flow.}$$

Price Prediction

- Empirical observation (Cont et al. '13):

$$\mathrm{d}s_t^{b/a} \approx \pm\delta \frac{\mathrm{OF}_{b/a}(t)}{D_{b/a}(t)}, \qquad D \dots \text{depths}, \quad \mathrm{OF} \dots \text{order flow.}$$

- First order approx for tick size $\delta = \mathrm{d}p$,

$$D_{a/b}(t) = \pm \int_0^{\delta} u_t(\pm p)\, \mathrm{d}p \approx \frac{\delta}{2} \nabla u_t(0\pm) = \frac{\delta\pi}{2L} Z_t^{a/b}, \tag{4}$$

$$\mathrm{OF}_{b/a}(t) \approx \mathrm{d}D_{b/a}(t) \tag{5}$$

After reparametrization:

$$\mathrm{d}D_{a/b}(t) = \nu_{a/b}(\mu_{a/b} - D_{a/b}(t))\, \mathrm{d}t + \sigma_{a/b} D_{a/b}(t)\, \mathrm{d}W_t^{a/b},$$

Price Dynamics

Induced Price Model I

$$\mathrm{d}s_t = \frac{1}{2}\left(\mathrm{d}s_t^b + \mathrm{d}s_t^a\right) = c_s\delta\left(\frac{\mathrm{d}Z_t^b}{Z_t^b} - \frac{\mathrm{d}Z_t^a}{Z_t^a}\right),$$

where $c_s \approx 1/2$, Z^a and Z^b are the factor processes from before.

Price Dynamics

Induced Price Model I

$$ds_t = \frac{1}{2}\left(ds_t^b + ds_t^a\right) = c_s\delta\left(\frac{dZ_t^b}{Z_t^b} - \frac{dZ_t^a}{Z_t^a}\right),$$

where $c_s \approx 1/2$, Z^a and Z^b are the factor processes from before.

When $Z^a = \mathcal{E}(\sigma_a W^a)$ and $Z^b = \mathcal{E}(\sigma_b W^b)$:

- Bachelier model with coefficients from order flow!
- $\langle s \rangle_t = \sigma_s^2 t$, where

$$\sigma_s = c_s\delta\sqrt{\sigma_a^2 + \sigma_b^2 - 2\varrho_{a,b}\sigma_a\sigma_b}.$$

Price Dynamics

Induced Price Model II

Summarizing the price dynamics are

$$\mathrm{d}s_t = c_s\delta\left(\frac{\mu_b}{D_b(t)} - \frac{\mu_a}{D_a(t)} + (\nu_b - \nu_a)\right)\mathrm{d}t + \sigma_b\,\mathrm{d}W_t^b - \sigma_a\,\mathrm{d}W_t^a,$$

$$\mathrm{d}D_{a/b}(t) = \nu_{a/b}(\mu_{a/b} - D_{a/b}(t))\,\mathrm{d}t + \sigma_{a/b}D_{a/b}(t)\,\mathrm{d}W_t^{a/b}$$

Price Dynamics

Induced Price Model II

Summarizing the price dynamics are

$$\mathrm{d}s_t = c_s\delta\left(\frac{\mu_b}{D_b(t)} - \frac{\mu_a}{D_a(t)} + (\nu_b - \nu_a)\right)\mathrm{d}t + \sigma_b\,\mathrm{d}W_t^b - \sigma_a\,\mathrm{d}W_t^a,$$

$$\mathrm{d}D_{a/b}(t) = \nu_{a/b}(\mu_{a/b} - D_{a/b}(t))\,\mathrm{d}t + \sigma_{a/b}D_{a/b}(t)\,\mathrm{d}W_t^{a/b}$$

- Time inhom. extension of classical Bachelier model, based on empirical observations!
- D_b and D_a are the market depth at bid and ask side.

Calibration: Setup

Model for depths:

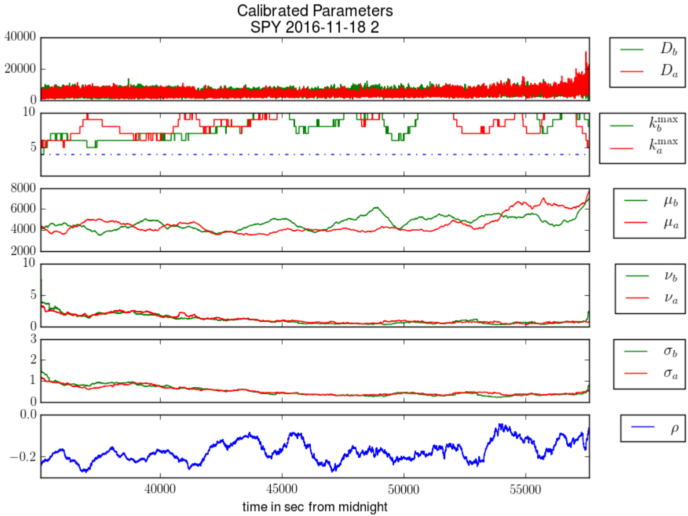
$$\mathrm{d}D_{a/b}(t) = \nu_{a/b}(\mu_{a/b} - D_{a/b}(t))\,\mathrm{d}t + \sigma_{a/b}D_{a/b}(t)\,\mathrm{d}W_t^{a/b}, \tag{6}$$

with $[W^a, W^b]_t = \rho t$. First try:

1. Split trading day in 10ms time intervals
   $\rightarrow$ observations $D_a(t_i)$, $D_b(t_i)$ of depths in first two levels, $i = 1, ..., N$.
2. For t_i estimate parameters based on $D_a(t_j)$ and $D_b(t_j)$,
   $t_j \in [t_i - 30\text{min}, t_i]$.
3. Recalibrate at $t_{i+5} = t_i + 50ms$.

Run on NASDAQ data for some of most liquid large-tick stocks and SPY+QQQ.

SPY 2016-11-16, depth (av in first 2 levels):



Price variations

Recall that for $c_s \approx \frac{1}{2}$,

$$\langle s \rangle_t = \sigma_s^2 t, \qquad \sigma_s = c_s \delta \sqrt{\sigma_a^2 + \sigma_b^2 - 2\varrho_{a,b}\sigma_a\sigma_b}.$$

$\longrightarrow$ Absolute (!) price variation determined by order flow volatility

Price variations

Recall that for $c_s \approx \frac{1}{2}$,

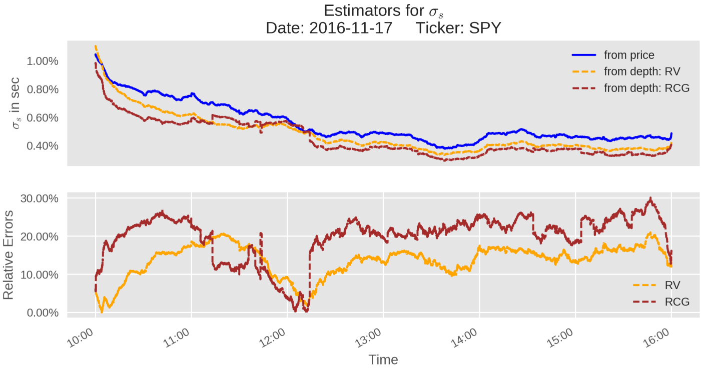
$$\langle s \rangle_t = \sigma_s^2 t, \qquad \sigma_s = c_s \delta \sqrt{\sigma_a^2 + \sigma_b^2 - 2\varrho_{a,b}\sigma_a\sigma_b}.$$

$\longrightarrow$ Absolute (!) price variation determined by order flow volatility

- Allows to check on market data if price model makes sense

Price variations

SPY 2016-11-17



С днем рождения

Bon anniversaire

Happy birthday

Buon compleanno

Herzlichen Glückwunsch zum Geburtstag

R. Cont, A. Kukanov, and S. Stoikov.
The price impact of order book events.
Journal of Financial Econometrics, 12(1):47–88, 2014.

J. Donier, J. Bonart, I. Mastromatteo, and J.-P. Bouchaud.
A fully consistent, minimal model for non-linear market impact.
Quantitative Finance, 15(7):1109–1121, 2015.

M. S. Müller.
A stochastic Stefan-type problem under first-order boundary conditions.
Ann. Appl. Probab., 28(4):2335–2369, 2018.

Symmetric Case

Assume $\mu_a = \mu_b = \mu$ and $\nu_a = \nu_b = \nu$, then

$$\mathbb{P}\left[s_t > s_0 + \frac{\delta}{2}\right] \approx \mathbb{P}\left[N < \frac{\nu\sqrt{t}}{2\sigma_s}\frac{\mu(D_0^a - D_0^b)}{D_0^b D_0^a} - \frac{1}{2\sigma_s\sqrt{t}}\right]$$

for a standard normal RV N.

Symmetric Case

Assume $\mu_a = \mu_b = \mu$ and $\nu_a = \nu_b = \nu$, then

$$\mathbb{P}\left[s_t > s_0 + \frac{\delta}{2}\right] \approx \mathbb{P}\left[N < \frac{\nu\sqrt{t}}{2\sigma_s}\frac{\mu(D_0^a - D_0^b)}{D_0^b D_0^a} - \frac{1}{2\sigma_s\sqrt{t}}\right]$$

for a standard normal RV N.

$\rightarrow$ Invers to Volume Imbalance VI = $D_0^b - D_0^a$!

VI as Price Predictor

- Accepted prediction method in academia and industry (e. g. Lipton et al (2013)):

$$\text{next price move} \sim \text{VI}$$

VI as Price Predictor

- Accepted prediction method in academia and industry (e. g. Lipton et al (2013)):

$$\text{next price move} \sim \text{VI}$$

- Cartea et al. (2015): MO triggers LO (on small time horizons - up to 100ms)
- However: Mean reversion yields expected OFI might be inverse to initial VI

VI as Price Predictor

- Accepted prediction method in academia and industry (e. g. Lipton et al (2013)):

$$\text{next price move } \sim \text{VI}$$

- Cartea et al. (2015): MO triggers LO (on small time horizons - up to 100ms)
- However: Mean reversion yields expected OFI might be inverse to initial VI
- Taking care of the time scales! (next price move vs scales of mean reversion)