Optimal Order Routing[1]

Jing Guo

Strats Associate, Goldman Sachs

PhD in Financial Engineering, Columbia University

March 4, 2018

[1] Reference: Maglaras (2015), C. Maglaras (n.d.)

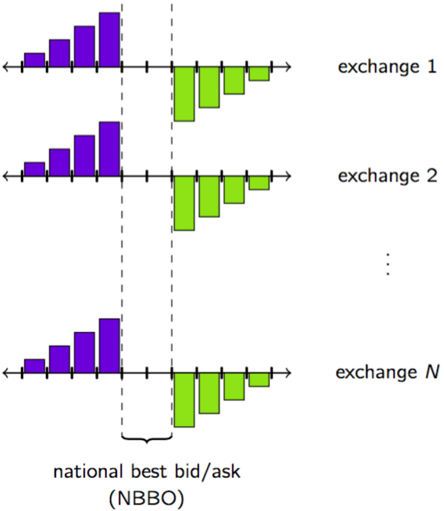
Outline

1. Order Routing in Fragmented LOB Markets
2. Order Routing Decision Problem
3. Data and Model Estimation

Algo Trading Systems: Typically Decomposed into 3 Steps

- **Trade scheduling (macro-trader):** splits parent order into $\sim$ 5 min slices (**Lecture 2**)
  - Relevant time-scale: minutes-hours
  - Schedule follows user selected strategy (VWAP, POV, IS, ...)
  - Reflects urgency, alpha, risk/return tradeoff
  - Schedule updated during execution to reflect price/liquidity/...
- **Optimal execution of a slice (micro-trader):** further divides slice into child orders (**Lecture 3**)
  - Relevant time-scale: secondsminutes
  - Strategy optimizes pricing and placing of orders in the LOB
  - Execution adjusts to speed of LOB dynamics, price momentum, ...
- **Order routing:** decides where to send each child order (**Lecture 4**)
  - Relevant time-scale: $\sim 1 - 50$ ms
  - Optimizes fee/rebate tradeoff, liquidity/price, latency, etc

Multiple Limit Order Books



- Price levels are coupled through protection mechanisms (Reg NMS[2]).

[2] *Regulation National Market System*

Time Scales

- Three relevant time scales:
  - **Events**: order/ trade/ cancellation interarrival times ($\sim$ ms - secs)
  - **Delays**: waiting times at different exchanges ($\sim$ sec - mins)
  - **Rates**: time-of-day variation of flow characteristics ($\sim$ min - hrs) [3]

---

[3]

- Arrival rate λ can be interpreted as, during small time interval Δt, the expected number of incoming orders is $\lambda\Delta t$.
- Arrival speed can be considered constant during this time range.

DJIA 30: Summary Statistics - Sept 2011

	Symbol	Listing Exchange	Price Low ($)	Price High ($)	Average Bid-Ask Spread ($)	Volatility (daily)	Average Daily Volume (shares, $\times 10^6$)
Alcoa	AA	NYSE	9.56	12.88	0.010	2.2%	27.8
American Express	AXP	NYSE	44.87	50.53	0.014	1.9%	8.6
Boeing	BA	NYSE	57.53	67.73	0.017	1.8%	5.9
Bank of America	BAC	NYSE	6.00	8.18	0.010	3.0%	258.8
Caterpillar	CAT	NYSE	72.60	92.83	0.029	2.3%	11.0
Cisco	CSCO	NASDAQ	14.96	16.84	0.010	1.7%	64.5
Chevron	CVX	NYSE	88.56	100.58	0.018	1.7%	11.1
DuPont	DD	NYSE	39.94	48.86	0.011	1.7%	10.2
Disney	DIS	NYSE	29.05	34.33	0.010	1.6%	13.3
General Electric	GE	NYSE	14.72	16.45	0.010	1.9%	84.6
Home Depot	HD	NYSE	31.08	35.33	0.010	1.6%	13.4
Hewlett-Packard	HPQ	NYSE	21.50	26.46	0.010	2.2%	32.5
IBM	IBM	NYSE	158.76	180.91	0.060	1.5%	6.6
Intel	INTC	NASDAQ	19.16	22.98	0.010	1.5%	63.6
Johnson & Johnson	JNJ	NYSE	61.00	66.14	0.011	1.2%	12.6
JPMorgan	JPM	NYSE	28.53	37.82	0.010	2.2%	49.1
Kraft	KFT	NYSE	32.70	35.52	0.010	1.1%	10.9
Coca-Cola	KO	NYSE	66.62	71.77	0.011	1.1%	12.3
McDonalds	MCD	NYSE	83.65	91.09	0.014	1.2%	7.9
3M	MMM	NYSE	71.71	83.95	0.018	1.6%	5.5
Merck	MRK	NYSE	30.71	33.49	0.010	1.3%	17.6
Microsoft	MSFT	NASDAQ	24.60	27.50	0.010	1.5%	61.0
Pfizer	PFE	NYSE	17.30	19.15	0.010	1.5%	47.7
Procter & Gamble	PG	NYSE	60.30	64.70	0.011	1.0%	11.2
AT&T	T	NYSE	27.29	29.18	0.010	1.2%	37.6
Travelers	TRV	NYSE	46.64	51.54	0.013	1.6%	4.8
United Tech	UTX	NYSE	67.32	77.58	0.018	1.7%	6.2
Verizon	VZ	NYSE	34.65	37.39	0.010	1.2%	18.4
Wal-Mart	WMT	NYSE	49.94	53.55	0.010	1.1%	13.1
Exxon Mobil	XOM	NYSE	67.93	74.98	0.011	1.6%	26.2

Figure 1: Descriptive statistics for the 30 stocks over the 21 trading days of September 2011. All statistics except the average bid-ask spread were retrieved from Yahoo Finance; the average bid-ask spread is a time average computed from our TAQ data set. The daily volatility is computed from closing prices over the period in question.

DJIA 30: Summary Statistics - Sept 2011

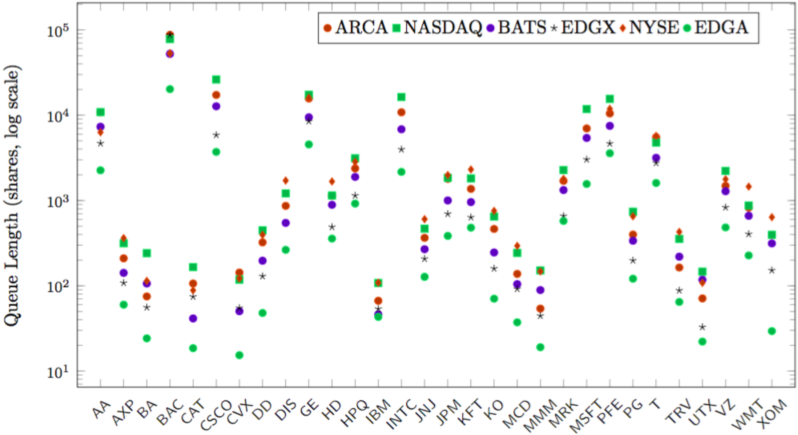


Figure 2: Average queue length (number of shares at the NBBO) across stocks and exchanges.

DJIA 30: Expected Delays - Sept 2011

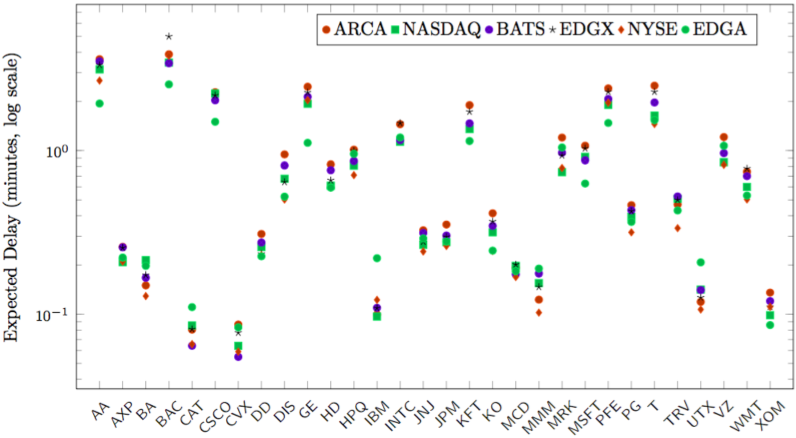
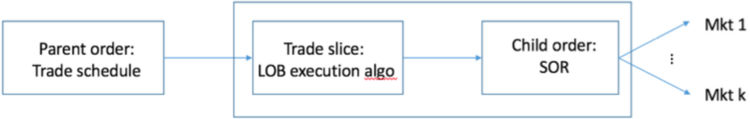


Figure 3: Average expected delay across stocks and exchanges.

- In general, expected delay is *negatively* related to queue length (liquidity).

Algo Engine Schematic: Smart Order Routing Problem



-
- What information is needed to optimize oder placement decision?
- What are the time scales of the trading decisions of the above modules?

Outline

1 Order Routing in Fragmented LOB Markets

2 Order Routing Decision Problem

3 Data and Model Estimation

The Limit Order Placement Decision [4]

- Factors affecting limit order placement:
  - Expected delay ($\approx$ 1 to 1000 seconds).
  - Rebates ($\approx$) \$-0.0001 a negative liquidity rebate is a fee charged to liquidity providers) to \$0.003 per share.
  - Rebates are significant in magnitude when compared to the bid-ask spread of a typical liquid stock of \$0.01 per share.
  - Other factors that affect decision such as short-term *alpha signals*, *estimates of adverse selection*, tiering agreements with exchanges.
- Mathematical formulation:

$$\tilde{r}_i := \text{rebate}_i + (\text{other factors}) = \text{"effective rebate"}$$
$$\mathbf{ED}_i := \text{expected delay.}$$

[4] C. Maglaras (n.d.)

Expected Delay

- Definition of *expected delay*

$$\mathbf{ED}_i := \frac{Q_i(t)}{\mu_i},$$

  where $Q_i(t)$ is the queue length at exchange i and μ_i is the market order arrival rate at exchange i.

- Traders choose to route their order to exchange i given by

$$\text{argmax}_{i \in \{1,2,\ldots,N\}} \quad \gamma \tilde{r}_i - \mathbf{ED}_i,$$

  where γ is a trade-off coefficient between price and delay, with units of time per dollar, that characterizes the type of the heterogeneous investors.

Attraction Model

- Attraction model

$$\mu_i(Q) := \mu \frac{f_i(Q_i)}{\sum_{j=1}^{N} f_j(Q_j)}.$$

  - $f_i(Q_i)$ captures attraction" of exchange i.
  - The larger queueing length Q_i, the higher rebate r_i, correspondingly the higher arrival market rate μ_i.

- C. Maglaras (n.d.) uses the simplified version:

$$f_i(Q_i) := \beta_i Q_i.$$

(Assume $\beta_i \sim 1/r_i$.)

The Market Order Placement Decision

- Market orders execute immediately, no queueing or adverse selection.
- Market orders incur fees ($\approx r_i$).
- Natural criterion is to route order according to

$$\text{argmin}_{i\in\{1,2,\ldots,N\}} \quad \{r_i : Q_i > 0\}$$

  where Q_i is queue length of exchange i (Only choose from those exchanges with positive queueing length).
- Routing decision differs from "fee minimization" due to:
  - Order sizes may have to be split across exchanges.
  - When Q_i is small, market order may not be filled completely.
  - Not all flow "optimized", may be under other economic considerations.
  - Traders avoid "clearing" queues to avoid increase price slippage.

Outline

1. Order Routing in Fragmented LOB Markets
2. Order Routing Decision Problem
3. Data and Model Estimation

Empirical Results

- NYSE TAQ data, millisecond timestamps
- Stocks: DJIA 30 Sept 2011
- 6 main exchanges (95% of lit" volume)
- Analysis uses time-averaged 60 min slices from 9:45am - 3:45pm $\times$ 20 days

	Exchange Code	Rebate ($ per share $\times 10^{-4}$)	Fee ($ per share, $\times 10^{-4}$)
BATS	Z	27.0	28.0
DirectEdege X (EDGX)	K	23.0	30.0
NYSE ARCA	P	21.0	30.0
NASDAQ OMX	T	20.0	30.0
NYSE	N	17.0	21.0
DirectEdege A (EDGA)	J	5.0	6.0

Table 1: Rebates and fees of the 6 major U.S. stock exchanges during the September 2011, per share traded.

Estimation of market order routing model (β's)

- Reduced form attraction" model for market order arrival rates:

$$\mu_i^{(s,j)}(t) := \mu^{(s,j)}(t) \frac{\beta_i^{(j)} Q_i^{(s,j)}(t)}{\sum_{k=1}^{N} \beta_k^{(j)} Q_k^{(s,j)}(t)},$$

  where s is the order size, j is the j'th security, i is the i'th exchange.
- Estimation procedure:
  - Measure $\mu_i^{(s,j)}(t)$, $\mu^{(s,j)}(t)$, $Q_i^{(s,j)}(t)$ by empirical frequency.
  - estimate $\beta_i^{(j)}$ via non-linear regression.

DJIA 30: Market order routing model (β's) Sept 2011

	Attraction Coefficient					
	ARCA	NASDAQ	BATS	EDGX	NYSE	EDGA
Alcoa	0.73	0.87	0.76	0.81	1.00	1.33
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.03)
American Express	1.19	1.08	0.99	0.94	1.00	0.94
	(0.02)	(0.02)	(0.04)	(0.03)	(0.00)	(0.06)
Boeing	0.95	0.67	0.81	0.74	1.00	0.73
	(0.02)	(0.01)	(0.01)	(0.02)	(0.00)	(0.04)
Bank of America	0.94	1.04	1.01	0.77	1.00	1.43
	(0.01)	(0.02)	(0.02)	(0.01)	(0.00)	(0.04)
Caterpillar	0.82	0.78	1.13	0.70	1.00	0.58
	(0.01)	(0.01)	(0.03)	(0.02)	(0.00)	(0.04)
Cisco	0.95	1.00	1.06	0.98	-	1.45
	(0.01)	(0.00)	(0.01)	(0.02)	-	(0.03)
Chevron	0.70	0.93	1.17	0.65	1.00	0.75
	(0.01)	(0.01)	(0.02)	(0.01)	(0.00)	(0.05)
DuPont	0.90	0.98	0.98	1.03	1.00	1.00
	(0.01)	(0.01)	(0.02)	(0.02)	(0.00)	(0.06)
Disney	0.69	0.88	0.78	0.88	1.00	1.04
	(0.01)	(0.01)	(0.02)	(0.03)	(0.00)	(0.03)
General Electric	0.79	1.01	0.94	0.73	1.00	1.63
	(0.01)	(0.01)	(0.02)	(0.01)	(0.00)	(0.03)
Home Depot	0.76	0.98	0.79	0.84	1.00	1.02
	(0.01)	(0.01)	(0.01)	(0.02)	(0.00)	(0.03)
Hewlett-Packard	1.04	1.04	1.02	0.68	1.00	0.82
	(0.02)	(0.01)	(0.02)	(0.02)	(0.00)	(0.03)
IBM	1.25	1.20	1.20	1.05	1.00	0.54
	(0.02)	(0.02)	(0.03)	(0.02)	(0.00)	(0.02)
Intel	0.83	1.00	0.96	0.84	-	1.04
	(0.01)	(0.00)	(0.01)	(0.02)	-	(0.03)
Johnson & Johnson	0.80	0.94	0.86	0.92	1.00	0.77
	(0.01)	(0.01)	(0.01)	(0.02)	(0.00)	(0.03)
JPMorgan	0.78	0.99	0.93	0.84	1.00	0.91
	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.02)
Kraft	0.72	0.89	0.83	0.73	1.00	1.06
	(0.01)	(0.01)	(0.01)	(0.02)	(0.00)	(0.03)
Coca-Cola	0.68	0.84	0.79	0.76	1.00	0.88
	(0.01)	(0.01)	(0.02)	(0.02)	(0.00)	(0.05)

DJIA 30: Pseudo R Square Sept 2011

	R_*^2		R_*^2		R_*^2
Alcoa	75%	Home Depot	87%	Merck	78%
American Express	64%	Hewlett-Packard	77%	Microsoft	80%
Boeing	75%	IBM	63%	Pfizer	79%
Bank of America	80%	Intel	82%	Procter & Gamble	80%
Caterpillar	58%	Johnson & Johnson	83%	AT&T	77%
Cisco	87%	JPMorgan	88%	Travelers	67%
Chevron	67%	Kraft	79%	United Tech	47%
DuPont	82%	Coca-Cola	81%	Verizon	79%
Disney	78%	McDonalds	74%	Wal-Mart	85%
General Electric	82%	3M	62%	Exxon Mobil	81%

- Definition of pseudo R_*^2:

$$R_*^2 := \frac{\text{Var}(||\text{ED}^{(s,j)}(t) - \tilde{\text{ED}}^{(s,j)}(t)||)}{\text{Var}(||\text{ED}^{(s,j)}(t)||)}$$

  - $\text{Var}(\cdot)$ = sample variance, averaged over all time slots t and both sides of market.
  - R_*^2 close to 1.

References I

C. Maglaras, C. C. Moallemi, H. Z. (n.d.). Queueing dynamics and state space collapse in fragmented limit order book markets.

Maglaras, C. (2015). Limit order book markets: a queueing systems perspective. https://www0.gsb.columbia.edu/faculty/cmaglaras/papers/IC-Lectures-2015.pdf.