© Théo Jalabert

Simulating and analyzing order book data: The queue-reactive model

Mathieu Rosenbaum

École Polytechnique

Fall 2020

© Théo Jalabert

Outline

© Théo Jalabert

Outline

© Théo Jalabert

Summary

Aim of this work

- Understanding the behaviours of market participants at different limits of the order book.
- Providing a realistic market simulator, enabling to compute execution costs of complex trading strategies.

Approach

- State (order book) dependent order flow intensities, in contrast to the Poisson approach.
- Empirical validation through full order book data analysis.

© Théo Jalabert

Modelling order book dynamics

Order book models

- Zero intelligence (Poisson flows) : Smith et al ; Cont, Stoikov and Talreja ; Cont and De Larrard ; Abergel and Jedidi.
- Hawkes flows : Hewlett, Large.
- Fokker-Planck dynamics with state dependence : Garèche, Disdier, Kockelkoren and Bouchaud.
- Equilibrium models : Rosu ; Lachapelle, Lasry, Lehalle and Lions.
- Functional approaches : Lakner, Reed and Stoikov ; Horst, Cont.

© Théo Jalabert

Modelling order book dynamics

Some missing elements in available models

- Market participants intelligence.
- Differences between traders behaviours at different limits.
- Dynamic bid-ask spread.
- Explaining the empirical shape of the order book.

© Théo Jalabert

Our approach

We consider the two following steps :

- First we model the dynamics of the order book for periods where a *reference price* stays constant.
- Then we introduce the dynamics of the reference price.

Our model

- No individual agent.
- Intelligence is added through a mean field game type approach.
- Empirical studies of traders average behaviours become possible.

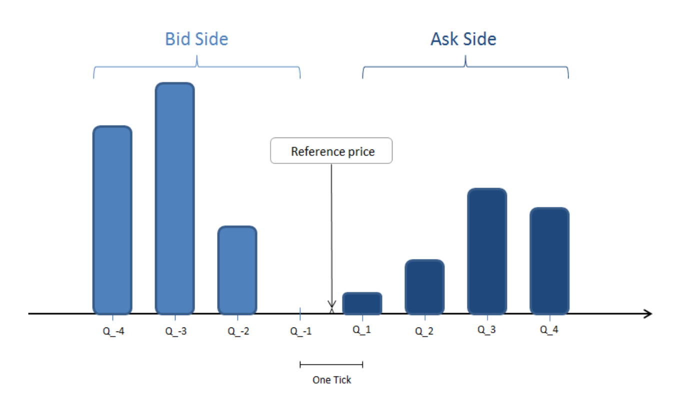
© Théo Jalabert

Outline

© Théo Jalabert

General framework

Limit order book in this framework



© Théo Jalabert

General framework

Assumptions

- p_{ref} stays constant.
- K limits are considered on each side.
- At the bid side of the order book, market participants can only send buy limit orders, cancel existing buy orders or send sell market orders.
- At the ask side of the order book, market participants can only send sell limit orders, cancel existing sell orders or send buy market orders.
- A constant order size is assumed for each limit Q_i (different values at different limits are allowed).

© Théo Jalabert

General framework

Limit order book as a continuous time Markov jump process

- The $2K$ dimensional LOB state $X(t)$: $X(t) = (Q_{-K}(t), \ldots, Q_{-1}(t), Q_1(t), \ldots, Q_K(t))$.
- Order flow intensities : $\lambda^M_{buy/sell}(x)$ (market orders), $\lambda^L_i(x)$ (limit orders at Q_i) and $\lambda^C_i(x)$ (cancellations at Q_i).
- The associated infinitesimal generator matrix $Q_{x,y}$ is

$$
\begin{aligned}
Q_{x,x+e_i} &= \lambda^L_i(x) \\
Q_{x,x-e_i} &= \lambda^C_i(x) + \lambda^M_{buy}(x) 1_{bestask(x)=i}, \text{if } i > 0 \\
Q_{x,x-e_i} &= \lambda^C_i(x) + \lambda^M_{sell}(x) 1_{bestbid(x)=i}, \text{if } i < 0 \\
Q_{x,x} &= -\sum_{y \in \Omega, y \neq x} Q_{x,y} \\
Q_{x,y} &= 0, \text{otherwise.}
\end{aligned}
$$

© Théo Jalabert

Theorem

Under reasonable technical assumptions, the $2K$ dimensional Markov jump process X is ergodic, which means that there exists a probability measure π that satisfies $\pi P(t) = \pi$ (π is called invariant measure, and $P_{xy}(t)$ is the transition probability from state x to state y in a time t) and for every x and y :

$$\lim_{t\to\infty} P_{xy}(t) = \pi_y.$$

© Théo Jalabert

Data description

Large tick stocks

- A large tick asset is defined as an asset whose bid-ask spread is almost always equal to one tick.
- Two French large tick stocks (France Telecom, now known as Orange, and Alcatel-Lucent) are studied in this work from Jan 2010 to March 2012. They exhibit very similar results.
- The stock France Telecom is chosen as illustration example in this presentation.

stock	number of orders per day	number of trades per day	spread size (ticks)
France Telecom	159250	7282	1.43
Alcatel Lucent	129400	8626	1.99

© Théo Jalabert

Estimating p_{ref}

Estimation method

- When the spread is odd (in tick unit) :

$$p_{ref} = p_{mid} = \frac{(p_{bestbid} + p_{bestask})}{2}.$$

- When it is even, we use either

$$p_{mid} + \frac{\text{tick size}}{2} \text{ or } p_{mid} - \frac{\text{tick size}}{2},$$

  choosing the one which is closer to the previous value of p_{ref}.
- More complex methods can be used for the estimation of p_{ref}. For large tick assets, remark that the estimated generator matrix remains stable under various methods of estimating p_{ref} (since it is essentially p_{mid}).

© Théo Jalabert

Model I : Collection of independent queues

Assumption

- Order flow arrival rates are functions of the target queue size only.

  $\rightarrow$ Birth and death process for each queue.

© Théo Jalabert

Empirical study : Collection of independent queues

Arrival/departure ratio $\rho_i(n)$

The order arrival/departure ratio $\rho_i(n)$, defined by

$$\rho_i(n) = \frac{\lambda_i^L(n)}{(\lambda_i^C(n+1) + \lambda_i^M(n+1))}.$$

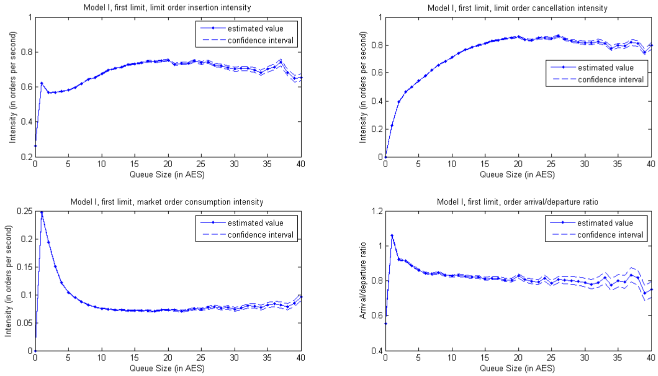
plays an important role in the stationarity and long term behaviour of the order book. We have that :

- The queue size tends to increase when $\rho > 1$.
- The queue size tends to decrease when $\rho < 1$.

© Théo Jalabert

Statistical properties of the order flow : first limits

Intensities as functions of the target queue size : first limits



© Théo Jalabert

Statistical properties of the order flow : first limits

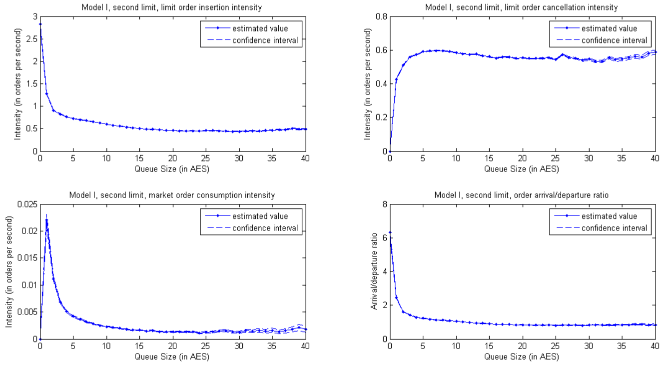
Comments

- Limit order insertion : almost a constant function, with a particularly smaller value observed at $Q_1 = 0$: limit order insertion within the bid-ask spread is risky.
- Cancellation : approximately an increasing concave function, due to the priority value of limit orders.
- Market order insertion : decreases exponentially with the available quantity at Q_1 : rushing for liquidity when it is rare and waiting for better price when liquidity is abundant.

© Théo Jalabert

Statistical properties of the order flow

Intensities as functions of the target queue size : second limits



© Théo Jalabert

Statistical properties of the order flow : second limits

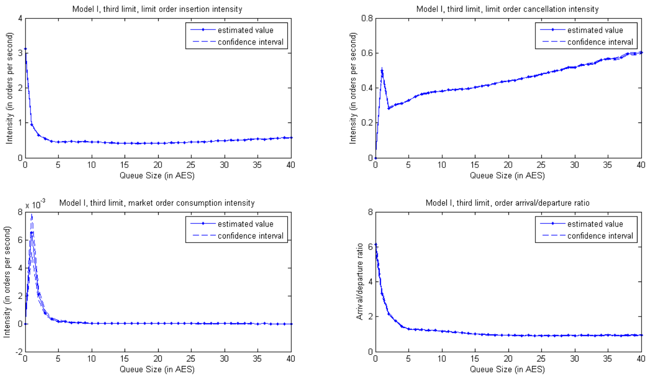
Comments

- Limit order insertion : a decreasing function of the queue size. A common strategy used in practice : posting orders at non-best limits when the queue size is small to seize priority.
- Cancellation : increases more rapidly, attains its asymptotic limit for a queue size around 5 AES (average event size).
- Market order insertion : market orders can arrive at Q_2 only when $Q_1 = 0$. The shape is very similar to that of Q_1 but the values are much smaller.

© Théo Jalabert

Statistical properties of the order flow

Intensities as functions of the target queue size : third limits



© Théo Jalabert

Statistical properties of the order flow : third limits

Comments

- Limit order insertion : remains a decreasing function as for the case of Q_2. Here market participants seem to stop rushing for future priority when the size becomes larger than 5 AES.
- Cancellation : now increases almost linearly as the queue size grows.
- Market order insertion : in some rare cases, one can still find market orders arriving at Q_3 (cross market orders or market orders occurring when the spread is large).

© Théo Jalabert

Asymptotic shape of the LOB

Stationary distribution

Denoting by $\pi_i(n)$ the stationary distribution of the queue size Q_i, the following result for the invariant distribution is easily obtained :

$$\begin{aligned}\pi_i(n) &= \pi_i(0)\prod_{j=1}^{n}\rho_i(j-1)\\ \pi_i(0) &= \Big(1+\sum_{n=1}^{\infty}\prod_{j=1}^{n}\rho_i(j-1)\Big)^{-1}.\end{aligned}$$

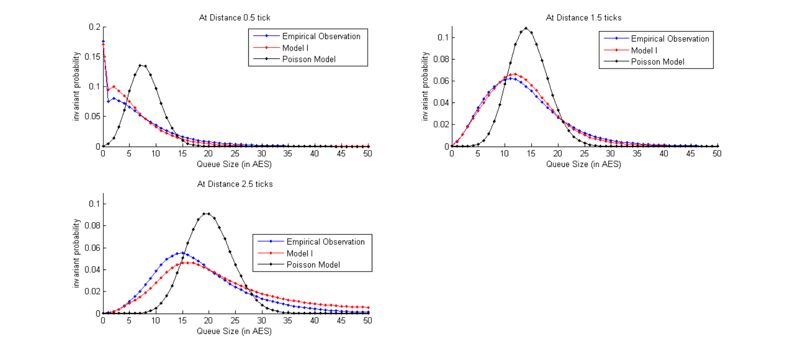
Remark

The long term behaviour of the order book is completely determined by the order arrival/departure ratio vector $\rho_i(n)$.

© Théo Jalabert

Asymptotic form of the LOB

Invariant distribution vs empirical distribution



© Théo Jalabert

Model II : Two sets of dependent queues

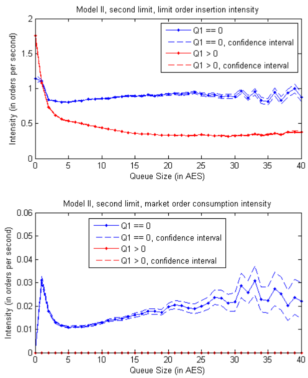
Assumptions

- Market orders are only sent to first and second limits.
- Market orders consume quantities at best limits.
- Dynamics at second limit depend also on whether the first limit is empty.

© Théo Jalabert

Statistical properties of the order flow

Intensities at Q2



© Théo Jalabert

Statistical properties of the order flow

Comments

- Limit order insertion : both curves are decreasing functions of the queue size, however, when Q_2 is the best ask limit, the order arrival rate reaches a much higher asymptotic value. The shape of the graph when $Q_1 > 0$ is again related to the arbitrage strategy mentioned previously.
- Cancellation : the cancellation rate is higher when $Q_1 = 0$.
- Market order : no market order can arrive at Q_2 when there are still quantities available at Q_1. Note that we treat cross limit market orders (that consume several limits) as a sequence of market orders arriving within a very short time period in our data.

© Théo Jalabert

Model II as a level independent quasi birth and death process

Assumption

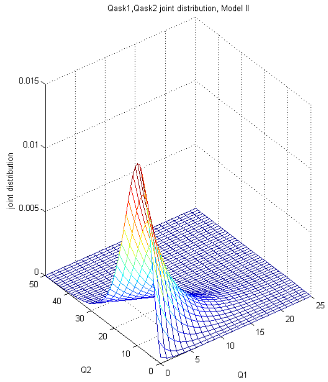
(Independent Poisson flows at Q_1) There are two positive constants λ_1 and μ_1, with $\lambda_1 < \mu_1$, such that for $k \geq 1$:

$$\begin{aligned} \lambda_1^C(k) + \lambda_{buy}^M(k) &= \mu_1 \\ \lambda_1^L(k) &= \lambda_1 \\ \lambda_1^L(0) &= \lambda_1. \end{aligned}$$

© Théo Jalabert

Asymptotic behaviours in Model II

Joint distribution of Q_1, Q_2, Model II



© Théo Jalabert

Model III : Modelling bid-ask dependences

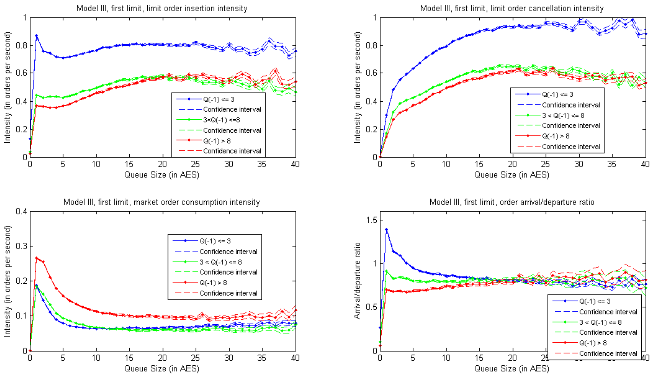
Assumptions

- Market participants adjust their trading rate not only according to the the target queue size, but also to whether the opposite queue size is small, usual or large.
- The regime switching in Model II still applies at $Q_{\pm 2}$.
- The problem is reduced to the study of the 4-dimensional process $(Q_{-2}, Q_{-1}, Q_1, Q_2)$.

© Théo Jalabert

Empirical study : Modelling bid-ask dependences

Intensity functions at Q1



© Théo Jalabert

Empirical study : Modelling bid-ask dependences

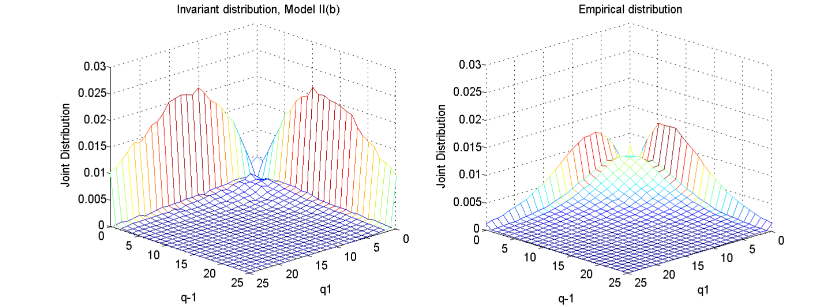
Comments

- Limit order insertion : a decreasing function of the opposite limit size.
- Cancellation : similar forms but different asymptotic values. It is not surprising that the cancellation rate, being an indicator of market participants patience, is a decreasing function of the liquidity level at the opposite side.
- Market order : when the volume available at Q_{-1} is abundant, more market orders are sent to Q_1. The reason is that in that case, transactions at Q_1 are relatively cheap as the fair price is temporarily closer to the price of Q_1.

© Théo Jalabert

Asymptotic behaviours in Model III, by Monte-Carlo method

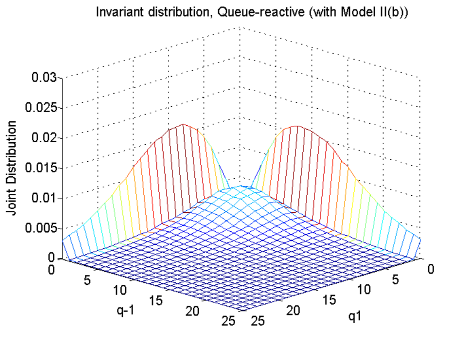
Joint distribution of Q_{-1}, Q_1



© Théo Jalabert

Asymptotic behaviours in Model III with dynamic reference price, by Monte-Carlo method

Joint distribution of Q_{-1}, Q_1



© Théo Jalabert

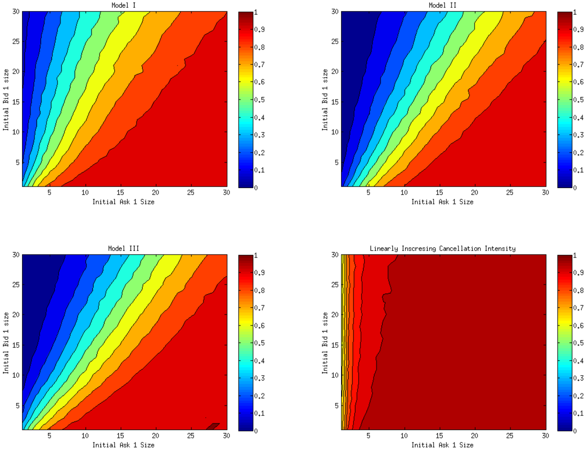
Example of application : Probability of execution

Problem

At $t = 0$, a trader posts a buy limit order at Q_{-1} and waits until either the order is executed or the opposite queue Q_1 is totally depleted. We want to estimate the probability of executing such order under different initial states of the order book.

© Théo Jalabert

Execution probability of a buying order placed at Q_{-1} at $t = 0$ (Monte-Carlo method)



© Théo Jalabert

Outline

© Théo Jalabert

Purely order book driven model

Dynamics of p_{ref} in the purely order book driven model

- We consider times of mid price move for changes in p_{ref}. More precisely, they are triggered with probability θ by one of the three following events :
  - The insertion of a buy (sell) limit order within the bid-ask spread, while Q_1 (Q_{-1}) is empty at the moment of this insertion.
  - A cancellation of the last limit order at the best offer queue.
  - A market order that consumes all the available quantity at the best offer queue.
- Q_i becomes either Q_{i+1} or Q_{i-1} depending on the direction of price move.

© Théo Jalabert

Purely order book driven model

Remarks

- Price fluctuations are completely generated by the order book dynamics.
- The price volatility is naturally an increasing function of θ.
- Strong mean-reverting behaviour of the price process.
- The maximum achievable volatility (mechanical volatility) is often smaller than the empirical volatility.

© Théo Jalabert

The queue-reactive model

Dynamics of p_{ref} in the queue-reactive model

- Changes of p_{ref} are triggered by changes in the mid price, with probability θ.
- Q_i becomes either Q_{i+1} or Q_{i-1} when p_{ref} changes.
- With probability θ^{reinit}, the whole LOB is redrawn from its invariant distribution around the new p_{ref} when p_{ref} changes.

Remarks

- We consider that, with probability θ^{reinit}, changes of price are due to exogenous informations, in which case market participants adjust very quickly their order flows around the new p_{ref}, as if a new LOB is redrawn from its invariant distribution.

© Théo Jalabert

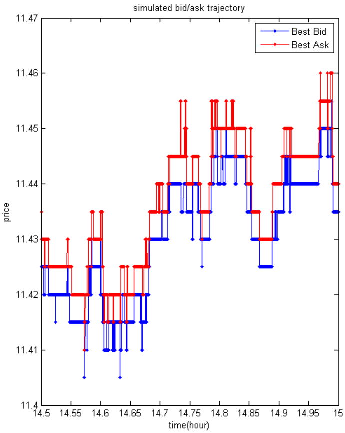
The queue-reactive model

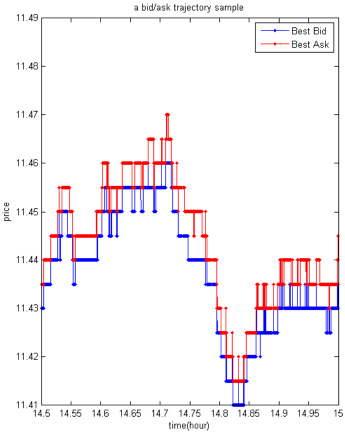
Calibration of θ and θ^{reinit}

- 10-minute price volatility of the asset.
- Mean-reversion ratio η of the mid price.

© Théo Jalabert

A simulated bid-ask trajectory





© Théo Jalabert

Theorem

After suitable rescaling, we obtain a diffusive behaviour at large time scales for the price in the queue reactive model.

© Théo Jalabert

Example of application : order placement analysis

The general optimal execution framework

- Trading horizon is divided into small slices (5-10 minutes).
- An execution algorithm determines, at the start of each slice, the quantity to be executed in that slice.

Order scheduling problem

- Almgren-Chriss approach.
- Optimal solution depends notably on the market impact function and the risk aversion ratio.

© Théo Jalabert

Example of application : order placement analysis

Order placement problem

- In each of these slices, how should the algorithm place orders ?
- Microstructural version of the volume scheduling problem, but is much more difficult to solve.
  - Price dynamics are no longer of Brownian type.
  - Queue priority starts to play an important role, as well as other microstructural features (tick size, order book state, trading speed).

© Théo Jalabert

Example of application : order placement analysis

Our approach

The queue-reactive model can be used as a market simulator for analyzing order placement tactics.

© Théo Jalabert

Example of application : order placement analysis

Two simple order placement tactics

- T1 : **"Fire and forget"** At $t = 0$, post a limit order at the best offer queue. When the mid price changes, cancel the limit order and send a market order at the opposite side with all the remaining quantities if any. At $t = T$, send all the remaining quantities at the opposite side to finish the execution.
- T2 : **"Pegging to the best".** At $t = 0$, post a limit order at the best offer queue, and then "peg" to it : if the best offer price changes, cancel the existing order and repost all the remaining quantities at the new best offer queue. If our order is the only remaining order in the best offer queue, cancel it and repost the remaining quantities at the newly revealed best offer queue. At $t = T$, send all the remaining quantities at the opposite side to finish the execution.

© Théo Jalabert

Example of application : order placement analysis

Order scheduling strategy

These two order placement tactics will be used together with two volume scheduling strategies (n_i denotes quantity to be trade at the i-th slice, and M is the number of slices.) :

- $n_{\text{total}} = 60$ AES, length of the subtrading periods : 10 minutes, $M = 20$.
- S1 : a linear scheduling $n_i = n_{\text{total}}/M$, VWAP benchmark.
- S2 : an exponential scheduling $n_i = n_{\text{total}}(e^{-(i-1)/4} - e^{-i/4})$, S_0 benchmark.

© Théo Jalabert

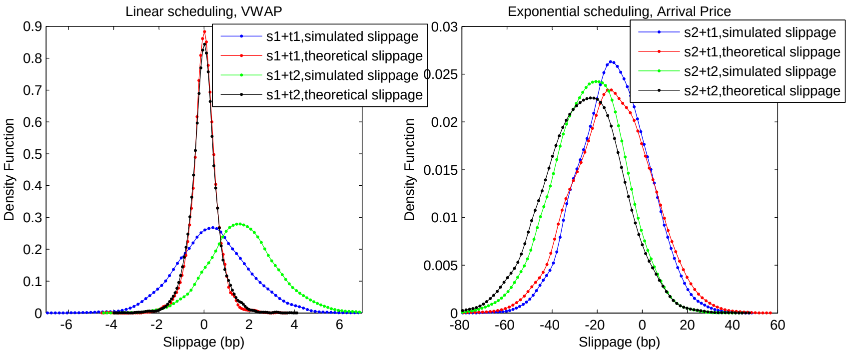
Example of application : order placement analysis

Decomposition of slippage

$$\begin{aligned}
\text{Slippage} &= \frac{P_{benchmark} - P_{exec}}{P_{benchmark}} \\
P^{theo}_{exec} &= \sum_{i=1}^{M} n_i \mathrm{VWAP}^i \\
\text{Slippage}^{theo} &= \frac{P_{benchmark} - P^{theo}_{exec}}{P_{benchmark}}.
\end{aligned}$$

© Théo Jalabert

Simulation results



© Théo Jalabert

Example of application : order placement analysis

Comments

- The tactic "Pegging to the best" performs better than the tactic "Fire and forget" for a VWAP execution benchmark. High passive execution rate, better average price than market VWAP, but large impact.
- The tactic "Fire and forget" slightly outperforms the tactic "Pegging to the best" for a S_0 execution benchmark. Lower passive execution rate but smaller impact.

© Théo Jalabert

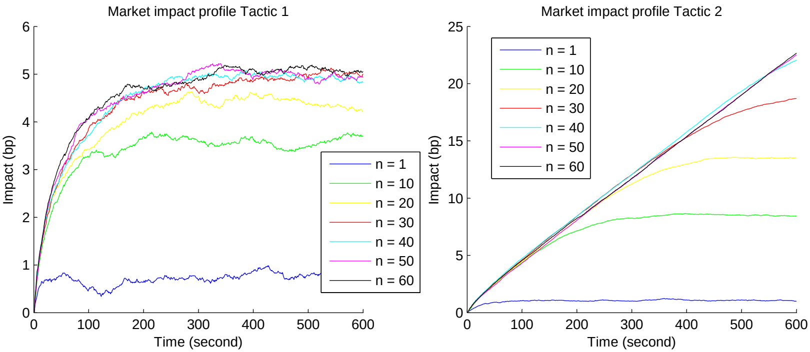
Example of application : order placement analysis

Market impact profiles

- An order placement tactic has two parameters : the period length T and the quantity to execute n. In the following experiments, T will be set to 10 minutes, and we vary the value of n from 1 to 60 AES.
- We denote by $\mathrm{MI}^i(t, n)$ the market impact at time t of Tactic i with target quantity n, defined by : $\mathrm{MI}^i(t, n) = E[\frac{S_t - S_0}{S_0}]$.

© Théo Jalabert

Market impact profiles of these two tactics



© Théo Jalabert

Market impact profiles

Comments

- The market impact curves are concave both in time and volume in our simulations.
- The price impact of the tactic "Fire and forget" is quite instantaneous and depends essentially on the target quantity n.
- The price impact of the tactic "Pegging to the best" is a more progressive process, depending both on the target quantity n and the duration t.