# The price impact of trades

**Julius Bonart**

Imperial College London, CFM-Imperial Institute of Quantitative Finance

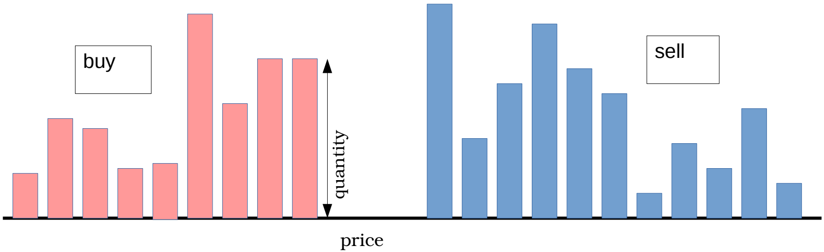
4 mars 2015

# Summary

# *Limit order books*

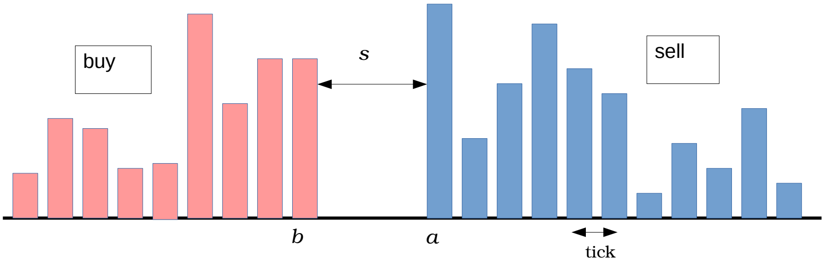
# The limit order book

Virtually all modern financial markets are based on *limit order books* which record buy or sell orders with specified quantities and prices.
*Transactions* take place when limit orders cross : Market order (MO)

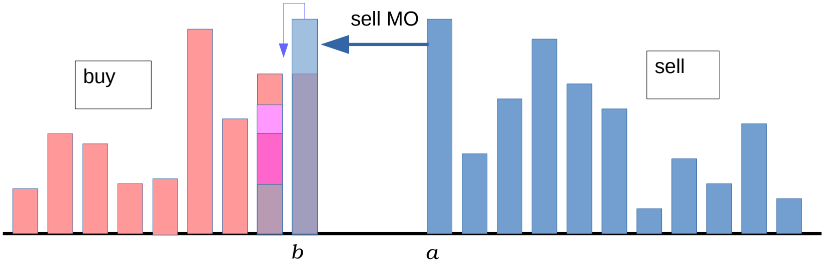


## The limit order book

Virtually all modern financial markets are based on *limit order books* which record buy or sell orders with specified quantities and prices.
*Transactions* take place when limit orders cross : Market order (MO)



## Transactions




Example : Sell limit order crosses bid which triggers a transaction. The sell MO eats up the buy LO according to *price priority* followed by *time priority* (of LO at the same price).

## Events in the order book

- New limit sell order
- New limit buy order
- Market sell order
- Maket buy order
- Cancellation of existing limit orders

Each market participant can choose between these actions on virtually all security exchanges.
Sometimes additional events exist : Hidden limit orders etc.

## Events in the order book

- New limit sell order
- New limit buy order
- Market sell order
- Maket buy order
- Cancellation of existing limit orders

Each market participant can choose between these actions on virtually all security exchanges. Sometimes additional events exist : Hidden limit orders etc.

The market reacts to all of these events ! ***Impact*** of limit order book events.

## Market microstructure

Theory and consequences of the impact of order book events : Determinants of the bid-ask spread, statistical arbitrage, etc. I will give a short illustration of some of these aspects. *Visible part of the limit order book*.

## Market "mesoscopic" microstructure

Behaviour of the market when large volumes are bought or sold. "Coarse-grained" approaches. Study of underlying *true* supply and demand curves : *Hidden (and much bigger) "part" of the limit order book*.

## Market microstructure

Theory and consequences of the impact of order book events : Determinants of the bid-ask spread, statistical arbitrage, etc. I will give a short illustration of some of these aspects. *Visible part of the limit order book*.

## Market "mesoscopic" microstructure

Behaviour of the market when large volumes are bought or sold. "Coarse-grained" approaches. Study of underlying *true* supply and demand curves : *Hidden (and much bigger) "part" of the limit order book*.

**Provocative statement**

The publicly visible limit order book does not reflect the true supply and demand of a financial market !

*Some important aspects of microstructure*

## The liquidity paradox

Liquidity providers submit limit orders. They *want to trade*.

But, liquidity providers do *not want to disclose their private information !*

Hence, liquidity providers *do not display* their buying/selling intentions.

## The liquidity paradox

Liquidity providers submit limit orders. They *want to trade*.

But, liquidity providers do *not want to disclose their private information !*

Hence, liquidity providers *do not display* their buying/selling intentions.

**Still not convinced ?**

The *total displayed volume* in a order book of a liquid stock is about $\sim 0.1\%$ of the daily total traded volume ! If you want to buy, for instance, 1.2% of a large company, there is *no way* you can achieve this at once !

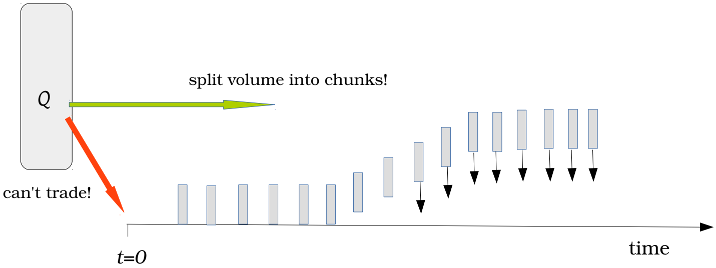
## The meta-order

So how do people trade large quantities ?

## The meta-order

So how do people trade large quantities ?

A meta-order is the *ensemble* of trades, executed *incrementally*, that belong to one single trading decision $\{\mathrm{buy/sell}, Q\}$.



## Consequence 1 : correlated order sign flow

Most trades belong to some meta-order. Hence, their ***sign*** is correlated with the signs of past and future trades !

Define $\epsilon(t) = \pm 1$ for buy/sell market order at time $t$. Then ***emprically***

$$\mathbb{E}[\epsilon(t)\epsilon(t+\tau)] \sim \tau^{-0.5} \ .$$

This is a ***long memory*** process and not in conflict with price efficiency (observed price is an almost perfect martingale).

Buys tend to be followed by buys and sells tend to be followed by sells.

## Consequence 2 : The bid-ask spread

Submit a *random* single BUY market order of volume $q$. It executes sell limit orders up to volume $q$,

- thereby the ask is *mechanically* impacted (moves up).
- hereafter the ask relaxes (on average !) and mean-reverts to the initial price. The mid-price is virtually not impacted.

## Consequence 2 : The bid-ask spread

Submit a *random* single BUY market order of volume $q$. It executes sell limit orders up to volume $q$,

- thereby the ask is *mechanically* impacted (moves up).
- hereafter the ask relaxes (on average !) and mean-reverts to the initial price. The mid-price is virtually not impacted.

But : Observed trades are immersed in a correlated order flow ! Observed impact of trade takes into account impacts of all future correlated trades.

## Consequence 2 : The bid-ask spread

But : Observed trades are immersed in a correlated order flow ! Observed impact of trade takes into account impacts of all future correlated trades.

Observe a single market buy order of volume $q$. It executes sell limit orders up to volume $q$,

- thereby the ask is *mechanically* impacted (moves up).
- hereafter the ask does *not* relaxe (on average !) ; the mid-price moves further up until reaching a plateau. One observes :

$$R = \sup_{t>0} \mathbb{E}\left[p(t) - p(0) \mid Buy\ MO\ at\ t = 0\right] > 0\ .$$

## Consequence 2 : The bid-ask spread

$$R = \sup_{t>0} \mathbb{E}\left[p(t) - p(0) \mid Buy\ MO\ at\ t = 0\right] > 0 \; .$$

Apply *competitive equilibrium methods* for the liquidity provider. Conditioned on the fact that his/her limit order is executes, he/she makes average P&L of $\mathcal{P} = \langle s \rangle/2 - \langle R \rangle$ :

$$\langle s \rangle/2 - \langle R \rangle \to 0^+ \; .$$

The relation $\langle s \rangle \approx 2\langle R \rangle$ is well satisified on liquid financial markets !

*Mesoscopic microstructure*

## Back to our meta-order

Such a meta-order is defined by its ***execution/trading rate*** :

$$m(t) = \frac{\partial}{\partial t}\mathrm{inventory}(t) \ .$$

Its ***volume*** :

$$Q = \int_{t_-}^{t_+} \mathrm{d}t \ m(t) \ ,$$

the ***starting and ending times***

$$t_- = \inf \mathrm{supp}(m) \ \ , \quad t_+ = \sup \mathrm{supp}(m) \ .$$

$T = t_+ - t_-$ is its ***length***.

## The impact costs

$$\mathcal{C} = \int_{t_-}^{t_+} \mathrm{d}t \; m(t)\mathbb{E}[p(t)] \; .$$

## The impact costs

$$\mathcal{C} = \int_{t_-}^{t_+} \mathrm{d}t \; m(t)\mathbb{E}[p(t)] \; .$$

We need to know how the price $p(t)$ reacts to the meta-order !

## The impact costs

$$\mathcal{C} = \int_{t_-}^{t_+} \mathrm{d}t \; m(t)\mathbb{E}[p(t)] \; .$$

We need to know how the price $p(t)$ reacts to the meta-order !

Introduce :

- peak impact $\mathbb{E}[p(t_+) - p(t_-)]$ of meta-order.
- impact trajectory $\mathcal{I}(t) = \mathbb{E}[p(t + t_-) - p(t_-)]$.

## Empirical formula for the peak impact

- peak impact $\mathcal{I}(T) = E[p(t_+) - p(t_-)]$ of meta-order.
- impact trajectory $\mathcal{I}(t) = \mathbb{E}[p(t + t_-) - p(t_-)]$.

Empirically, we find for the peak impact :

$$\mathcal{I}(T) \approx Y\sigma\sqrt{\frac{Q}{ADV}} ,$$

with :

- $Y \sim 1$,
- $\sigma$ the daily volatility of the security,
- $ADV$ the average daily traded volume of the security.

## The square-root law

$$\mathcal{I}(T) \approx Y\sigma\sqrt{\frac{Q}{ADV}},$$

- Impact is *non-linear* in the volume $Q$.
- Impact is *non-Markovian* : Remember that the volume $Q$ is executed incrementally. Each single trade cannot have the same impact, otherwise total impact would be linear in $Q$. The first trade has larger impact than subsequent trades !
- Impact is approximately independent of the execution path and depends only on the total volume $Q$. A more detailed study does show variations with respect to the execution path : Optimal execution problem (later).

## Towards a theory of market impact : Remember the liquidity paradox

Remember : Liquidity providers hide their intentions, they offer far less liquidity as is needed to execute large volumes. The consequences of this behaviour can be directly observed on financial markets : Existence of meta-orders, correlated order flow, etc.

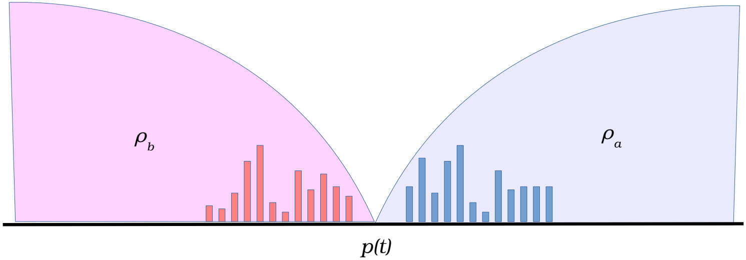
## Concept of the latent order book

Remember : Liquidity providers hide their intentions, they offer far less liquidity as is needed to execute large volumes. The consequences of this behaviour can be directly observed on financial markets : Existence of meta-orders, correlated order flow, etc.

True supply and demand is *not* displayed in the order book. There is a fictitious non-public book which records the intentions of market participants : The ***latent order book***.

## Concept of the latent order book

True supply and demand is *not* displayed in the order book. There is a fictitious non-public book which records the intentions of market participants : The ***latent order book***.



## Dynamics of the latent order book

- At each time step a fraction $\rho$ of all *latent* orders at each point $y$ moves with $\mathrm{d}W_t$,

$$\mathrm{d}W_t \mathrm{d}W_s = \sigma^2 \delta(t-s)\mathrm{d}t \ .$$

*Explanation : This models exogeneous information instantly digested by a part $\rho$ of the market participants.*

- Define the ask and bid densities of the latent order book : $\rho_a(y,t)$ and $\rho_b(y,t)$. Set $\varphi(y,t) = \rho_a(y,t) - \rho_b(y,t)$.
- Define $\phi(x,t) = \varphi(y,t)$ with $x = y - \rho \int^t \mathrm{d}W_s$.

*Explanation : This is a change of reference frame. We model the midprice $p(t) = \rho \int^t \mathrm{d}W_s$ as a Brownian motion, $\phi$ is then the latent order book observed in the co-moving reference frame of $p(t)$, i.e. $x$ measures the relative distance with respect to $p(t)$.*

## Dynamics of the latent order book

Define $\phi(x,t) = \varphi(y,t)$ with $x = y - \rho \int^t \mathrm{d}W_s$.

$$\dot{\varphi}(y,t) = -\rho \mathrm{d}W_t \varphi'(y,t) + \frac{\rho}{2}\sigma^2 \varphi''(y,t) \, ,$$

$$\dot{\varphi}(y,t) = \dot{\phi}(x,t) - \rho \mathrm{d}W_t \phi'(x,t) + \frac{\rho^2}{2}\sigma^2 \phi''(x,t) \, .$$

This yields

$$\dot{\phi}(x,t) = \frac{\sigma^2}{2}\rho(1-\rho)\phi''(x,t) \, .$$

Set

$$D = \frac{\sigma^2}{2}\rho(1-\rho) \, .$$

**$\phi$ satisfies a simple diffusion equation !**

## Market clearing in the latent order book

Note that : Market clearing can only occur in the real book !

Define : $R(x) = r\rho_a(x)\mathbf{1}_{x<0} + r\rho_b(x)\mathbf{1}_{x>0}$, with $r$ the reaction rate.

*Explanation : The reaction rate corresponds to the fraction of latent orders, on the wrong side of the mid-price, that is submitted as (real) limit or market orders at the mid-price per unit time. The number of executed orders per unit time must be equal to $J/D$ in the stationary case.*

## Market clearing in the latent order book

Define : $R(x) = r\rho_a(x)\mathbf{1}_{x<0} + r\rho_b(x)\mathbf{1}_{x>0}$, with $r$ the reaction rate.

*Explanation : The reaction rate corresponds to the fraction of latent orders, on the wrong side of the mid-price, that is submitted as (real) limit or market orders at the mid-price per unit time. The number of executed orders per unit time must be equal to $J/D$ in the stationary case.*

We have already defined $\phi(x) = \rho_a(x) - \rho_b(x)$ (with $x$ measured in the co-moving reference frame). We need to study the second superposition $\psi(x) = \rho_a(x) + \rho_b(x)$.

*Objective : The solution $\psi(x)$ depends on the inhomogeneity $R(x)$. Since $R(|x|) = r[\psi(x) - \phi(|x|)]$ we find a self-consistent equation for $R(x)$.*

## Market clearing in the latent order book

Solution :

$$R(x) \sim |x|\mathrm{K}_1(\sqrt{J/2D} \cdot |x|) \simeq \begin{cases} \text{const.} & |x| \to 0 \\ e^{-\sqrt{J/2D}\cdot|x|} & |x| \to \infty \end{cases}$$

The overlap has a finite width $\sqrt{D/J}$. The amplitude of the overlap scales with $1/r$.
**Hence : When $r \to \infty$ the overlap vanishes independently of the available liquidity $\mathcal{L} = J/D$.**

## Market clearing in the latent order book

The overlap has a finite width $D/J$. The amplitude of the overlap scales with $1/r$. **Hence : When $r \to \infty$ the overlap vanishes independently of the available liquidity $\mathcal{L} = J/D$.**

The market clearing mechanism thus works according to :

- The price movements lead to latent outstanding liquidity (latent buy orders above and latent sell orders below the current market price).
- These orders cannot be instantly executed due to restricted liquidity. Instead, traders disclose at each time step a fraction $r\mathrm{d}t$ of the outstanding orders and submit it at real market/limit orders.
- This mechanism leads to market clearing in the latent book : When $r$ is big, outstanding liquidity becomes small such that the incoming flow of new orders is kept constant. Hence, the limit $r \to \infty$ is commensurate with restricted liquidity in the real book.
- The current market price and the latent market clearing price coincide !

## Towards the price impact formula

Homogeneous problem with boundary conditions :

$$\partial_t \phi(x,t) = D\partial_{xx}\phi(x,t) \ ,$$
$$\lim_{x\to\pm\infty} \partial_x \phi(x,t) = J \ .$$

## Towards the price impact formula

Consider an additional trader who wishes to execute his metaorder. Non-homogeneous problem :

$$\partial_t \phi(x,t) - D\partial_{xx}\phi(x,t) = m(t)\delta(x - x_t) \ ,$$
$$\lim_{x\to\pm\infty} D\partial_x\phi(x,t) = J \ ,$$

where $m(t)$ is the trading rate and $x_t$ is the mid-price, defined via

$$\phi(x_t, t) = 0 \ .$$

(*market clearing*)

## The price impact formula

Solution :

$$\phi(x,t) = -\frac{J}{D}x + \int_{-\infty}^{\infty} \mathrm{d}y \int_0^t \mathrm{d}t' \frac{m(t')}{\sqrt{4D\pi(t-t')}} e^{-\frac{(x-y)^2}{4D(t-t')}} \delta(y - x_{t'}) \, .$$

Integrate over the Dirac-distribution and use that $x_t$ is a zero of $\phi$ :

$$x_t = \frac{D}{J} \int_0^t \mathrm{d}t' \frac{m(t')}{\sqrt{4D\pi(t-t')}} e^{-\frac{(x_t - x_{t'})^2}{4D(t-t')}} \, .$$

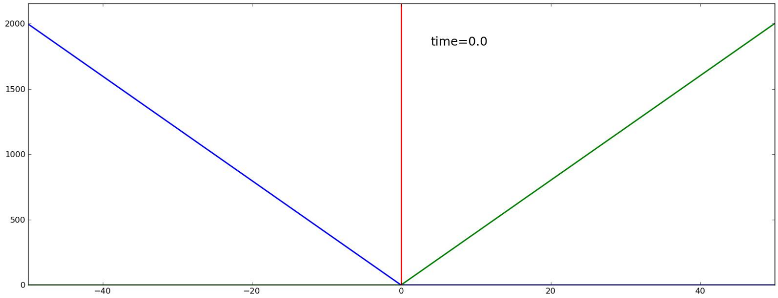
Implicit equation for $x_t$ which depends on the whole history $\{x_{t'<t}\}$.

## Example of impact

Constant $m$ and relaxation after meta-order is completed. With additional noise in the order book.

## Example of impact

Constant $m$ and relaxation after meta-order is completed. With additional noise in the order book.



## Absence of dynamical arbitrage

Define the execution cost :

$$\mathcal{C}[m] = \int_0^T \mathrm{d}t \; m(t) x_t \; .$$

Then we have the following **theorem** :

Let $x_t$ be the solution of our model for some *non trivial round-trip* $m$, i.e. $\int \mathrm{d}t \; m(t) = 0$ and $m \neq 0$. Then, execution costs are strictly positive, $\mathcal{C} > 0$.

Let $x_t$ be the solution of our model for some *non trivial round-trip* $m$, i.e. $\int \mathrm{d}t \; m(t) = 0$ and $m \neq 0$. Then, execution costs are strictly positive, $\mathcal{C} > 0$.

**One-line proof** :

Use the price impact formula of our model to replace $x_t$ :

$$\mathcal{C} = \frac{1}{2} \int_0^T \mathrm{d}t\mathrm{d}s \; m(t) M(t,s) m(s) \; ,$$

with

$$M(t,s) = \frac{D^2}{J} \int_{-\infty}^{\infty} \mathrm{d}z\mathrm{d}u \; K(t,u;z) K^*(s,u;z) \; ,$$
$$K(t,u;z) = z \cdot \mathbf{1}_{u \leq t} \cdot e^{-Dz^2(t-u) + izx_t} \; .$$

Hence $M$ is a square and thus positive.

## Square-root solution for constant trading rate

$$x_t = A\sqrt{Dt}$$

is an exact solution with

$$A = \frac{m}{J}\int_0^1 \mathrm{d}\eta \, \frac{e^{-\frac{A^2(1-\sqrt{\eta})}{4(1+\sqrt{\eta})}}}{\sqrt{4\pi(1-\eta)}} \, .$$

## Square-root solution for constant trading rate

$$A = \frac{m}{J} \int_0^1 \mathrm{d}\eta \, \frac{e^{-\frac{A^2(1-\sqrt{\eta})}{4(1+\sqrt{\eta})}}}{\sqrt{4\pi(1-\eta)}} \, .$$

is such that for large $m$ impact is square-root in volume since $A^2 \simeq 2m/J$ :

$$x_T \sim \sqrt{Q} \, .$$

## Square-root solution for constant trading rate

$$A = \frac{m}{J}\int_0^1 \mathrm{d}\eta \, \frac{e^{-\frac{A^2(1-\sqrt{\eta})}{4(1+\sqrt{\eta})}}}{\sqrt{4\pi(1-\eta)}} \, .$$

is such that for large $m$ impact is square-root in volume since $A^2 \simeq 2m/J$ :

$$x_T \sim \sqrt{Q} \, .$$

The model reproduces the empirically observed square-root law. Impact is square-root not only at its peak but throughout the whole impact trajectory.

## Beyond the quare-root law : Small constant trading rate

$$A = \frac{m}{J} \int_0^1 \mathrm{d}\eta \, \frac{e^{-\frac{A^2(1-\sqrt{\eta})}{4(1+\sqrt{\eta})}}}{\sqrt{4\pi(1-\eta)}} \, .$$

is such that for *small m* impact is with $A \simeq m/J$ :

$$x_T \sim \sqrt{mQ} \, .$$

The model predicts vanishing impact for small trading rates $m$ ! Not accounted for in the empirical formula.

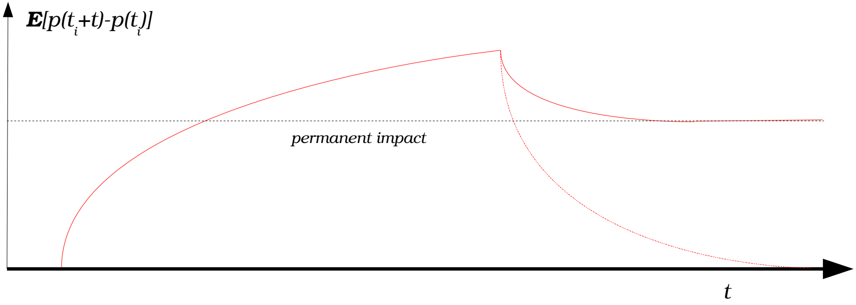
## (Non-)Mechanical impact

Our model predicts the mechanical impact, i.e. the impact of a meta-order *without* any information content (random meta-order).

If the meta-order bears information, an additional non-mechanical "impact" arises (not really an impact as not stimulated by the meta-order itself but by its conditioning to the direction of the market)

After completion of the meta-order, mechanical impact decreases to zero. However, non-mechanical impact remains positive if order bears information.

## Impact of “informed” and “random” meta-order



**Problem of finding the "bare" mechanical impact**

How to "subtract" the information from the empirically observed impact to obtain the mechanical impact ?

**Answer :**

Subtract permanent impact from observed peak impact.

*Empirical analysis on the Bitcoin/USD exchange market*

## Empirical analysis : What has been done ?

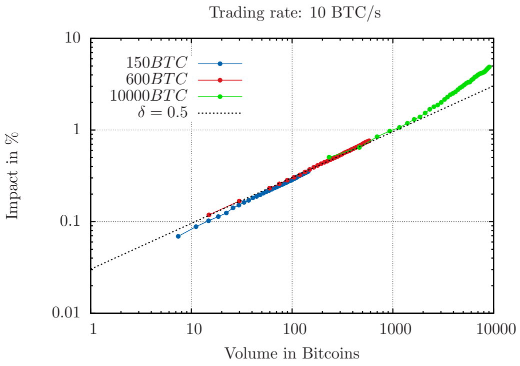
- Square-root law holds on all stock/future markets (the "empirical formula"). Permanent impact is $\approx 2/3$ of peak impact (2/3-law)
- No convincing analysis of market impact for small $m$. The "subtraction trick" has not been used in the past.
- Empirical analysis is difficult : Insider knowledge is necessary to reconstitute the meta-orders from anonymous trades.

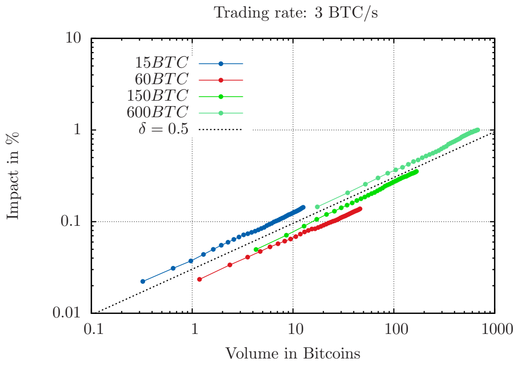
## Empirical analysis

*We have a complete dataset of 2 million meta-orders on the BitCoin/USD exchange market*

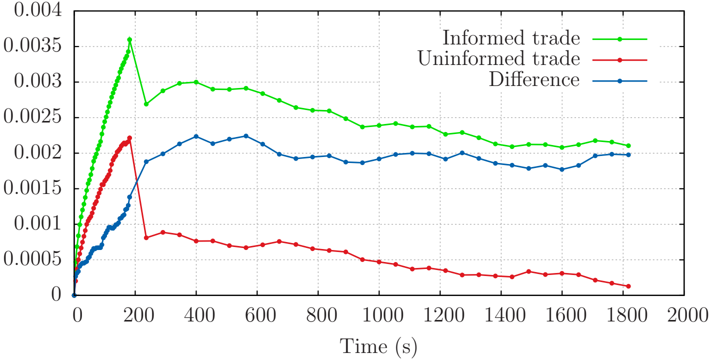
- Square-root law holds on the BitCoin, despite absence of sophisticated arbitrage ( !)
- Mechanical impact vanishes for small trading rate $m$ ! (no good picture, yet)

## Square-root law on the Bitcoin : even trajectory-wise !



## Impact of informed/uninformed meta-orders on the Bitcoin



*Summary of empirical evidence*

## Summary of empirical evidence : BitCoin, futures and stocks

- Impact is square-root : Peak and trajectory. Hence, impact is *non-linear* and *non-Markovian*.
- Subtraction trick shows that small trading rates lead to smaller impact at variance with the standard empirical formula and in agreement with theoretical model based on latent liquidity.
- Permanent impact is $\approx 2/3$ of peak impact.
- Uninformed trades (uncorrelated with the underlying oderflow ; cashflow trades) do not have any permanent impact. Hence, permament impact is a phenomenon solely based on the information content of the meta-order.

*Applications and further aspects*

**Application 1 : Optimal execution**

Liquidation horizon $T$, quantity $Q$ fixed. $\mathcal{F}$ space of possible execution rates $m : [0, T] \to \mathbb{R}$.

$$\mathcal{C} : \mathcal{F} \to \mathbb{R}^+$$
$$\arg\max_{m \in \mathcal{F}} \mathcal{C}[m] \ ?$$

*All examples in the literature treat optimal execution within a Markovian setup :*

$$\mathbb{E}[p_{t+\mathrm{d}t} | p_t] = \mathrm{const.} \cdot \mathrm{d}t \cdot f(m_t) \ .$$

Even if $f(m)$ is non-linear, above impact model does not correspond to reality !

*Generalization to non-Markovian impact : Ongoning work !*

## Application 2 : Stock pinning

*Empirical observation :* Spot price tends to converge to one of the possible strike prices on stocks with heavy derivative use such as Apple and Google.

*Assumption 1* : Option sellers (banks) $\Delta$-hedge, buyers (industry companies) do not (or less).
*Assumption 2* : Impact is *linear* : $\dot{x}_t = \mathcal{L}^{-1}\dot{Q}$ (*major contradiction with our work*).

Banks hold $-C_t + \Delta(t, x_t)x_t$ and

$$\mathcal{L}\dot{x}_t = \frac{\mathrm{d}}{\mathrm{d}t}\Delta(t, x_t) = \dot{\Delta} + \dot{x}_t\Delta'$$
$$\Rightarrow \dot{x}_t = \frac{\partial\Delta(t, x_t)/\partial t}{\mathcal{L} - \partial\Delta(t, x_t)/\partial x}\ .$$

$\dot{\Delta} \gtrless 0$ for $x_t \gtrless S$ and $\lim_{t\to T} \Delta'(t, S) \to \infty$. Hence, small enough $T - t$ there the strike is a local attractor !

*How to generalize to non-linear impact ?*

**Appendix 1 : Farmer-Lillo-Gerig-Waelbroeck-model and the $2/3$-law**

*Objective : Calcuate impact profile during the execution of a meta-order of length $T$*

$x_t$ is the execution price. Then

$$x_{t+1} = \begin{cases} x_t + s_t^+ \\ x_t - s_t^- \end{cases}$$

$P_t =$ Probability at $t$ that meta-order continues at $t$

$$= \mathbb{P}[T > t | T > t-1] = \frac{\mathbb{P}[T \geq t+1, T \geq t]}{\mathbb{P}[T \geq t]} = \frac{\mathbb{P}[T \geq t+1]}{\mathbb{P}[T \geq t]} .$$

## Farmer-Lillo-Gerig-Waelbroeck-model and the 2/3-law

*The martingale condition*

$$s_t^-(1-P_t) = s_t^+ P_t \ .$$

$$x_t = \mathbb{E}[x_\infty | t] = (1-P_t)\mathbb{E}[x_\infty | \text{MO stops at } t] + P_t(x_t + s_t^+) \ .$$

*The fair-pricing condition*

$$\mathbb{E}[x_\infty | \text{MO stops at } t] = \frac{1}{t}\sum_{k=0}^{t} x_k \ .$$

(All trades have unit volume).

## Farmer-Lillo-Gerig-Waelbroeck-model and the $2/3$-law

We find :

$$tx_t = ts_t^- + \sum_{k=0}^{t} x_k \ ,$$

and finally

$$\frac{t-1}{t}\frac{1-P_t}{1-P_{t-1}}s_{t-1}^+ = P_t s_t^+ \ .$$

Note that empirically $\mathbb{P}[T \geq t] \sim t^{-\gamma}$ with $\gamma \approx 3/2$. For large $t \gg 1$ we have

$$s_t^+ \simeq \left(1 - \frac{2-\gamma}{t}\right) s_{t-1}^+ \simeq t^{\gamma-2} s_0^+$$

$$x_t \sim t^{\gamma-1} \sim \sqrt{t}$$

$$\mathbb{E}[x_\infty | \mathrm{MO\ stops\ at}\ t] \simeq \frac{2}{3} x_t \ .$$

*Thank you !*