Market impact as anticipation of the order flow imbalance

Thibault Jaisson

Ecole Polytechnique, CMAP

Wednesday 5th March 2014
Outline

1. Introduction

2. Linking the order flow and the price
   - Market efficiency and linear permanent market impact
   - Market efficiency and the martingale hypothesis
   - Our model independent relation

3. Application to the Hawkes order flow example
   - Hawkes processes
   - Computation of a propagator model
   - Long memory in the order flow
   - Power law impact function
Outline

1. Introduction
2. Linking the order flow and the price
3. Application to the Hawkes order flow example
Basic concepts

- Limit orders, liquidity, order book.
- Market orders, order flow, long memory, power law, market impact.
- Metaorders, impact function, power law, transience, permanent impact.

Figure: Schematized impact function.
Theoretical studies:

Empirical studies:
- Bershova and Rakhlin 2013.
- Waelbroeck and Gomes 2013.
Apparent paradox persistence-market efficiency

A solution is proposed by Lillo et al. 2004 for the case of a discrete time FARIMA order flow, ie. such that the sign of market orders $\varepsilon$ checks

$$E[\varepsilon_t | F_{t-1}] = \sum_{i \geq 1} a_i \varepsilon_{t-i}$$

the price process should satisfy

$$p_t = p_{t-1} + \theta(\varepsilon_t - \sum_{i \geq 1} a_i \varepsilon_{t-i})$$

which is equivalent to Bouchaud et al. 2004:

$$p_t = p_0 + \sum_{k \geq 0} b_k \varepsilon_{t-k}$$

with $b_k = \theta(1 - \sum_{i=1}^{k} a_i)$. 

Thibault Jaisson
Outline

1. Introduction

2. Linking the order flow and the price
   - Market efficiency and linear permanent market impact
   - Market efficiency and the martingale hypothesis
   - Our model independent relation

3. Application to the Hawkes order flow example
There are two fundamental ways to understand why orders should impact the price:

- **Informational**: Orders reflect a private information that investors have on a stock value.
- **Mechanical**: Orders create an unbalance in supply and demand.
If the impact is mechanical, then the impact of trades does not depend on the reason why they are executed. If moreover there is no price manipulation opportunities\(^1\). Then, in various frameworks such as

- Huberman and Stanzl 2004,
- Gatheral 2010,

if there is permanent market impact then it is linear in volume.

\(^1\)A strategy that begins and ends with no inventory is not profitable on average.
The martingale hypothesis

- Usual prices (Ask, Bid, Mid, Last) are not martingales. They are often mean reverting at high frequency.
- In many works, it is assumed that there exists an underlying price which is a martingale. That is what we do here.
- In fact, under very weak hypotheses on the efficiency of the market dynamics, it is always possible to define a martingale fair price (ongoing work).
Our hypotheses

Assumption

*The permanent impact of metaorders is linear in their volume.*

Assumption

*The price is a martingale.*

Assumption

*The flow of metaorders is equal to the flow of market orders.*
Computation

We denote \((v^a_i)\) (resp. \((v^b_i)\)) the volumes of the \(N^a\) (resp. \(N^b\)) buy (resp. sell) metaorders of the trading day \([0, S]\) and \(V^a_t\) (resp. \(V^b_t\)) the cumulated volume of buy (resp. sell) market orders executed on \([0, t]\).

\[
\begin{align*}
    P_S &= P_0 + \kappa \left( \sum_{i=1}^{N^a} v^a_i - \sum_{i=1}^{N^b} v^b_i \right) + Z_S \\
    &= P_0 + \kappa (V^a_S - V^b_S) + Z_S,
\end{align*}
\]

where \(\kappa\) is a positive constant and \(Z\) is a martingale random walk. Therefore:

\[
\begin{align*}
P_t &= \mathbb{E}[P_S | \mathcal{F}_t] \\
    &= P_0 + \mathbb{E}[\kappa (V^a_S - V^b_S) | \mathcal{F}_t] + \mathbb{E}[Z_S | \mathcal{F}_t].
\end{align*}
\]
Considering that the trading day is long, we get the following result.

**Theorem**

The price process is equal to the anticipation of the order flow imbalance:

\[ P_t = P_0 + \kappa \lim_{s \to +\infty} \mathbb{E}[V_s^a - V_s^b | F_t] \]  

(1)

where \( V_s^a \) is the cumulated volume of ask orders executed on \([0, s]\).
The price is a martingale even if the order flow exhibits persistence.

The price process only depends on the global market order flow and not on the individual executions of metaorders.

This formula can be seen as a rigorous and model independent generalization of the postulate “the impact is proportional to the innovation in the order flow”.
1. Introduction

2. Linking the order flow and the price

3. Application to the Hawkes order flow example
   - Hawkes processes
   - Computation of a propagator model
   - Long memory in the order flow
   - Power law impact function
Definition of Hawkes processes

A Hawkes process $N$ is a point process on $\mathbb{R}$ of intensity:

$$
\lambda_t = \mu + \int_{-\infty}^{t} \phi(t - s) dN_s 
$$

(2)

$$
= \mu + \sum_{J_i < t} \phi(t - J_i)
$$

(3)

where $\mu \in \mathbb{R}^*$ is the exogenous intensity and $\phi$ is a positive kernel supported in $\mathbb{R}_+$ which satisfies $\int \phi < 1$ and the $J_i$ are the points of $N$. 
Basic properties

Proposition (Hawkes 1971)

The process is well defined and admits a version with stationary increments under the stability condition:

\[ |\phi| := \int \phi < 1. \]

Proposition

The average intensity of a stationary Hawkes process is

\[ \Lambda := E[\lambda_t] = \frac{\mu}{1 - |\phi|}. \]
Applications of Hawkes processes

Hawkes processes can be calibrated (see Ozaki 1979 or Dayri et al. 2012) to reproduce the clustering present in point processes. This explains their wide attractivity.

- Seismology, see Ogata 1998.
- Genomic analysis, see Reynaud-Bouret and Schbath 2010.
- Traffic network, see Brémaud and Massoulié 2002.

In finance they have already been used to model:

- Price changes, see Bacry and Muzy 2013 and Hardiman et al. 2013.
- Order book events, see Large 2007.
- Financial contagion, see Aït-Sahalia et al. 2010.
Cluster representation of Hawkes processes (I)

Let us consider the population model:

- At time 0 there are no individuals.
- Migrants arrive as a Poisson process of intensity $\mu$.
- If a migrant arrives in $t$, its children are the points of an inhomogeneous Poisson process of intensity $\phi(\cdot - t)$.
- If a child is born in $t$, its children are the points of an inhomogeneous Poisson process of intensity $\phi(\cdot - t)$.
- Etc...

Proposition (Hawkes, Oakes 1974)

The cumulated number $N$ of individuals who arrived and were born before $t$ is a point process of intensity:

$$\lambda_t = \mu + \sum_{J_i < t} \phi(t - J_i).$$
Cluster representation of Hawkes processes (II)

- If we call colony the set of descendants of a migrant, a Hawkes process can be characterized as a Poisson superposition of independent colonies.
- The average cardinal of a colony checks
  \[ \mathbb{E}[\# C] = 1 + |\phi| \mathbb{E}[\# C] \]
  and thus \( \mathbb{E}[\# C] = 1/(1 - |\phi|) \).
- \( 1 - |\phi| = \frac{\mu}{\Lambda} \) is the proportion of exogenous orders.
- \( |\phi| = \frac{\Lambda - \mu}{\Lambda} = \frac{\mathbb{E}[\# C] - 1}{\mathbb{E}[\# C]} \) is the proportion of endogenous orders.
All market orders are of volume $v$.

The number of market orders at the ask $N^a$ and at the bid $N^b$ are two independent Hawkes processes of intensity $\mu$ and kernel $\phi$. 
Propagator formula

**Proposition**

In the Hawkes order flow model, assuming that the price process checks Equation (1), then it follows a continuous time propagator model:

$$P_t = P_0 + \int_0^t \zeta(t-s)(dN_s^a - dN_s^b),$$  \hspace{1cm} (4)

with \( \zeta(t) = \kappa \nu \frac{1}{1-\int_{-\infty}^t \phi(s)ds} \left( (1 - \int \phi) + \int_t^{+\infty} \phi(s)ds \right). \)

The proof is a straightforward calculus using the Hawkes prediction formulas.
Necessary martingale condition on \( \zeta \) and \( \phi \)

We could have computed the \( \zeta \) more easily. Indeed, in the case of a continuous time propagator model combined with a Hawkes order flow,

\[
\frac{E[P_{t+h}|\mathcal{F}_t] - P_t}{h} \to \zeta(0) \int_0^t \phi(t-s)(dN^a_s - dN^b_s) + \int_0^t \zeta'(t-s)(dN^a_s - dN^b_s).
\]

Therefore:

**Proposition**

*If the price is a martingale, then:*

\[
\zeta'(x) = -\zeta(0)\phi(x).
\]

This is analogous to:

\[
b_k - b_{k-1} = -b_0a_k.
\]
Long memory and criticality

The infinitesimal covariance of a Hawkes process

$$\text{Cov}(|t - t'|) := \frac{\mathbb{E}[dN_t dN_{t'}] - \Lambda^2 dt dt'}{dt dt'}$$

can be linked in the Fourier domain to the kernel by

$$\hat{\text{Cov}}(z) = \frac{\mu}{1 - \int \phi} \frac{1}{|1 - \hat{\phi}(z)|^2},$$

we get that the long memory in the order flow,

$$\int \text{Cov} = \hat{\text{Cov}}(0) = +\infty,$$

implies the criticality of the Hawkes process, $|\phi| = \hat{\phi}(0) = 1$.

Such Hawkes processes are shown to exist in Brémaud and Massoulié 2001.
Nearly unstable Hawkes processes

- If we consider a critical Hawkes process with $|\phi| = 1$, then we must take $\mu = 0$ in order to have a stationary version of the process.
- This would lead to a completely endogenous market.
- Instead, we choose to consider a nearly unstable Hawkes process with $0 < 1 - |\phi| \ll 1$, see Jaisson and Rosenbaum 2013.
- We formally show that if $\phi(x) \sim \alpha c^\alpha/x^{1+\alpha}$, with $\alpha \in ]0, 1/2[$, then such a process exhibits long memory up to the time scale $c/(1 - |\phi|)^{1/\alpha}$ as it is suggested by Hardiman et al. 2013 and Bacry and Muzy 2013.
Introduction
Linking the order flow and the price
Application to the Hawkes order flow example

Hawkes processes
Computation of a propagator model
Long memory in the order flow
Power law impact function

Apparent long memory for nearly unstable Hawkes processes

More precisely:

**Proposition**

*For time scales $\tau$ such that $c \ll \tau \ll c(1 - |\phi|)^{-1/\alpha}$:

$$\text{Cov}(\tau) \simeq \frac{K}{|\tau|^{1-2\alpha}}.$$*
Hawkes processes, long memory and market endogeneity

We show that the endogeneity results can be very model dependent. Let us consider that the order flow is a superposition of independent metaorders somehow as in Lillo et al. 2004. If the size distribution of metaorders is

\[ P(i) = \frac{1}{(i-1+m)^{2+\gamma}} \left( \sum_{k \geq 1} \frac{1}{(k-1+m)^{2+\gamma}} \right) \]

with \( \mathbb{E}[\text{Length}] - 1 \sim m \ll 1 \) and \( \gamma \in ]0, 1[ \).

- On the one hand, since \( m \) is small metaorders are almost always composed of one market order (and \( \mathbb{E}[\text{Length}] \sim 1 \)) and the order flow is deeply exogenous therefore we want to take \( |\phi| \sim 0 \).

- On the other hand, the order flow presents long memory therefore estimations should yield \( |\phi| \sim 1 \).
In the above framework, $N^a$ and $N^b$ are the flows of anonymous market orders.

This corresponds to modelling the market from the point of view of a passive agent who does not use orders and does not know who uses the different orders.

In order to compute the impact function of metaorders, it is convenient to look at the market from the point of view of someone who is executing a (buy) metaorder.

To do that we model our order flow as a Poisson process $P^{F,T}$ of intensity $F$ on $[0, \tau]$. The total buy and sell order flows are thus $N^a + P^{F,T}$ and $N^b$. 
The market reacts in the same way to our order flow as to that of the rest of the market. Therefore, the price process during our metaorder checks

$$P_t = P_0 + \int_0^t \zeta(t - s)(dN^a_s - dN^b_s) + \int_0^t \zeta(t - s) dP_s^{F,T}.$$ 

Taking expectations yields

$$MI(t) := \mathbb{E}[P_t - P_0] = F \int_0^{t\wedge\tau} \zeta(t - s) ds.$$
Power law impact function for nearly unstable Hawkes processes

We go back to the framework of the power law nearly unstable Hawkes process. We want to compute the impact function of the metaorder after a time $\tau$ such that: $c \ll \tau \ll c(1 - |\phi|)^{-1/\alpha}$:

$$MI(\tau) = F \int_0^\tau \zeta(\tau - t)dt = F \int_0^\tau \zeta(t)dt$$

with

$$\zeta(t) = kv(1 - |\phi| + \int_t^{+\infty} \phi(s)ds) \approx \frac{K}{t^\alpha}$$

if $c \ll t \ll c(1 - |\phi|)^{-1/\alpha}$.

Therefore:

$$MI(\tau) \approx K\tau^{1-\alpha}.$$
For a nearly unstable Hawkes process whose kernel’s shape is asymptotically a power law of exponent $1 + \alpha$ with $\alpha \in ]0, 1/2[$:
- The long memory exponent is equal to $\gamma = 1 - 2\alpha$.
- The impact exponent is equal to $\nu = 1 - \alpha$.

Therefore, these two exponents are linked by: $\nu = (1 + \gamma)/2$. 

Thibault Jaisson
Market impact as anticipation of the OFI
If we look at the distribution queue of the infinitesimal correlation, we can easily estimate $\gamma$:

![Log/log infinitesimal covariance](image)

**Figure**: Log/log infinitesimal covariance (QuantHouse, xFDAX July 2013).

We get $\gamma \sim 0.2$ for the DAX which means that the impact exponent should be $\nu \sim 0.6$ which is consistent with the data.
Conclusion

- We have computed a general relation between the order flow and the price process which solves the “paradox”: Order flows are persistent and yet prices are martingales.
- We applied this relation to a simple Hawkes order flow.
- We retrieved a link between the long memory of the order flow and the concavity of the impact function.
Thank you for your attention!